

Reinforcing access to research data

The existence of public databases adds an obligation to some authors to deposit data. The situation with research materials is less clear-cut. It is time, nevertheless, for a change in *Nature's* policies.

WHAT does it mean to publish a scientific result? A paper in a journal is a necessary means of communicating results; sometimes, however, it is not considered sufficient. Occasional disputes have arisen among astronomers, for example, when some have been denied prompt access to satellite data. But the issue is most pervasive in molecular biology, where public access to data and to research materials is considered by many to be an obligatory adjunct to publication. Hitherto, *Nature* has not imposed on its authors the condition that they make data or materials available. That is about to change, though more selectively than some might wish.

It is now taken for granted by many who derive macromolecular structures using X-rays that the structural coordinates should be lodged in a publicly accessible database. There is also a move to ensure that structure amplitudes, from which coordinates are derived, should be registered (see page 202). Similar obligations to deposit data are felt by most who report gene sequences. For both structures and sequences, public databases are now well established, and easily accessed for both the registration and the supply of data.

From now on, *Nature* will hold to the principle that such information should be made available to the public as a condition of publication. No paper reporting such data submitted after the date of this issue will be accepted without a reference either to a database accession number, or to an Internet address at which the data can be accessed from the time of publication. In imposing those conditions (which apply equally to publicly and privately funded researchers), *Nature* will wish to be assured that the game is being played fairly: that registered structures, for example, will include all atoms rather than simply a carbon backbone. But if a community widely supports a database that allows access to be restricted for a specified period, as do the structural biologists, so be it.

Restrictions can affect access to research materials as well as to data. Many types of material can be affected, but none more so than 'knock-out mice'. Since the late 1980s, geneticists have been able to introduce DNA 'constructs' into cells and disrupt or replace targeted genes in mice. Each construct consists of a vector — which makes it possible to obtain large quantities of the DNA in question — into which the modified form of the gene can be integrated. The construct is generally inserted into certain lines of cultured embryonic cells, which can then be incorporated into an early embryo. Many biologists regard the resulting organisms as common property, whether produced with private or public funds: the mouse becomes part of the publication.

Although logically valid, this principle quickly leads to difficulties. One relates to return on investment: knock-out mice, especially, are expensive, demand months of work, and represent a resource that provides a significant competitive edge. Is it fair, the moment the first results are published, to insist that the mouse be made available to all? If so (and not all agree), funding agencies will need to impose far stricter

conditions and sanctions than obtain at present. But even if immediate distribution were to become the rule, the practical difficulties carry funding implications. There are costs and staff time involved in looking after and shipping the mice, and in ensuring that mice arriving in a laboratory carry no risks of contamination. Many centres established to distribute genetically modified organisms — *Drosophila* and the nematode worm *Caenorhabditis elegans* are notable examples — function efficiently. But with mice, for which the burden of tasks required to provide access is much greater, such centres are under increasing strain, and cannot now accept everything they are offered.

It is clear from a spot poll by *Nature* of this art's practitioners that these are controversial matters, and that attempts to improve the situation by, for example, asking authors to sign undertakings of good intent can prove futile. There is every reason to anticipate that, as publicly funded distribution centres and targeting technologies develop, access to materials will become less of a problem. For now, *Nature* intends to be pragmatically positive. Accounts of genetically altered animals will in future appear in our pages only on condition that the authors undertake to supply the DNA constructs used on request. In cases where the authors are prepared to make further materials (such as embryos or mice) available for non-commercial use, they will also be encouraged to state how these can be obtained — whether from a distribution centre, or an author's laboratory. The assumption will be that users of such materials will, in subsequent publications, give credit to originators by citing the appropriate paper. And those who flagrantly play the game unfairly may themselves become targets — of the written word. □

Enter *Nature Biotechnology*

The family of *Nature* journals is about to expand once again, to develop a publishing niche.

SINCE the launch in 1983 of *Bio/Technology* by Nature Publishing Co. Inc., a subsidiary of Macmillan Magazines Ltd, the publication has pursued a unique role as an outlet for papers across the spectrum of its discipline. Those papers have been accompanied by news and comment in a fashion that has much in common with *Nature* and its more recent sibling journals: *Nature Genetics*, *Nature Structural Biology*, and — founded a year ago — *Nature Medicine*.

Given the strong citation record of *Bio/Technology*, it makes good sense to transform it into a *Nature* journal. Accordingly the number of pages devoted to original research is to be doubled and the title adjusted. The aim will be to highlight new applications of technology to biology and vice versa, in the broadest way possible. Welcome, from March, *Nature Biotechnology*. □



Ocean drilling enters choppy waters as France and UK question strategy

Paris. France and the United Kingdom are considering withdrawing from the 20-year-old international Ocean Drilling Program (ODP). Critics in both countries claim that the scientific output of the programme no longer provides sufficient value for money, and that continued participation is difficult to justify, given the stiff competition for shrinking national research budgets.

The ODP studies the two-thirds of the Earth's crust that lies under water. Deep-ocean drilling — the ODP ship can drill cores in water up to 8.2 km deep — provides the means for carrying out research into the geological processes that create continents, for example, while the geological record also provides information on climate change.

The US National Science Foundation pays more than half of the \$45-million annual budget of ODP. The rest comes from its six other members: the United Kingdom, France, Germany and Japan, a consortium of smaller European countries organized through the European Science Foundation, and a joint membership of Canada and Russia.

France's participation is supervised by the marine research agency, IFREMER, and costs around FF40 million (US\$8.2 million) annually. Britain's involvement is run by the National Environmental Research Council (NERC), costing around £3 million (US\$4.6 million) a year. But both are now reviewing their continued participation in ODP beyond 1998, when the current agreement expires.

The French review, which has already been completed, recommends that France withdraw from the ODP. The panel of Earth scientists that carried out the review is particularly critical of the fact that most ODP results have been published as in-house 'grey' literature, claiming that only a few have given rise to papers in international peer-reviewed journals.

"The present scientific output does not justify the costs," says Vincent Courtillot, the rapporteur of the review, who is also head of the Laboratory of Palaeomagnetism and Geodynamics at the Institut de Physique du Globe de Paris. France's tight budget for Earth sciences could be spent better, he says, adding that ODP has run out of steam, and needs to be replaced by a new international programme with clearer scientific goals and strategies.

The scientific content of the ODP is coordinated through a body known as the Joint Oceanographic Institutions Deep

Earth Sampling (JOIDES). Half the subscriptions from members go towards paying for the operation of a converted oil exploration ship, the *JOIDES Resolution*. But the review claims that "ODP's [scientific] choices sometimes seem to have been aimed more at finding uses for [the ship], rather than addressing scientific problems at the lowest cost".

Scientific criteria have been sacrificed, claims the review, in doomed attempts by ODP to please two distinct groups of users with very different needs. Climatologists and

with domestic competition for funds from scientists not involved in ODP.

But echoes of some of the French criticisms are also being heard in the United Kingdom. NERC officials say that the questions are whether ODP has produced sufficient results for the money spent on it and whether it is continuing to produce results at the same rate as in its early years. "In these days of tight budgets we can't have a large 'ring-fenced' area in our budget that is immune from competition with other areas," says one.

^{ODP} The United Kingdom has traditionally been a strong supporter of ODP. Bids by its researchers have fared well within the programme, and the ODP's science planning office has been located at the University of Wales in Cardiff since 1993. But even some of the British scientists working in the programme agree that it needs to be focused more on major scientific problems.

The preliminary findings of the NERC review, chaired by Chris Hawksworth of the Open University, are expected to be known later this month. John Krebs, the chief executive of NERC, will then hold a meeting with Pierre David, the director general of IFREMER, to reach a common position. This meeting will be followed by one that includes Germany, traditionally a strong supporter of ODP.

Three options are likely to be on the table: continuing within the existing ODP, or withdrawing from ODP and either associating with a planned Japanese ocean drilling programme or establishing an autonomous European programme. Either of the last two possibilities might also be developed into a new global programme incorporating the US ship.

Indeed, Japan has already allocated funding to a feasibility study of a drilling ►

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Uncharted course: British and French reviews may lead to changes for the Ocean Drilling Program.

palaeoclimatologists tend to be interested in sediments and to prefer short-duration stays at many sites. But geologists tend to be interested in the deeper crust, and to prefer visits of longer duration at fewer sites. The only solution, argues the review, is for each user group to have its own ship, tools and research strategy.

The conclusions of the review have been hotly contested by ODP officials, who claim in private that French dissatisfaction with the programme stems from a resentment at both the US leadership of the programme and at having won fewer bids than the scientists concerned would have liked, combined

New director for Cambridge biology laboratory

London. The UK Medical Research Council's Laboratory of Molecular Biology (LMB) in Cambridge is to have a new director. Richard Henderson, currently head of the laboratory's structural studies division, will take over from Sir Aaron Klug, who has recently become the new president of the Royal Society.

Henderson's appointment is being seen as a sign that the LMB is keen to maintain its pre-eminence in the study of molecular structures. The Louis-Jeantet Prize for

Medicine — which Henderson won in 1993 for his work on the structure of bacteriorhodopsin — has been awarded this year to his colleague Nigel Unwin, joint head of the division (see page 200).

Henderson is due to take over as director of the LMB at the beginning of October. The timing of his appointment coincides with the arrival of George Radda, currently head of the department of biochemistry at the University of Oxford, as the new chief executive of the research council. □

► ship that would cost \$300-400 million and be twice the size of the *JOIDES Resolution*. A decision on whether to build the ship is expected this year.

But according to one NERC official, Britain is unlikely to be keen on restricting its involvement in the planned Japanese effort alone. Japan's interest in earthquakes, he argues, means that the planned Japanese ship could be tied up for years studying plate tectonics in oceanic areas off the Japanese coast. At the same time, the official adds that a European solution — being promoted by France — "is difficult to imagine from a British point of view, at present".

This could change if the European Union (EU) were to back such a programme, he says. But such an outcome seems improbable. The European Commission recently turned down a ECU7-million (US\$9.1 million) proposal for a project, CORSAIRES, involving shallow drilling in sediments. Instead it approved an ECU500,000 feasibility study for the project.

But beyond this, the commission has no interest in deep ocean drilling, according to one official from the Marine Science and Technology programme (MAST). The costs of establishing an ocean drilling programme would require greatly increased funding for the MAST programme, which currently has a budget of ECU228 million.

Moreover, the political will to support deep ocean drilling is lacking both within the top ranks of the commission and among most member states, according to the official.

One possible outcome of the tripartite meeting could be an agreement to continue with ODP until 2003 on the understanding that a new international programme would then be established that would include the Japanese ship. This option is said to be favoured by Germany, although one NERC official says that the prospect of the United Kingdom and France simply withdrawing from ODP "cannot be ruled out".

But such threats are interpreted by some as simply a tactic aimed by the United Kingdom and France at helping to achieve a reduction in their subscriptions to the programme. Scope for such reductions, they point out, might come from the expected enlargement of ODP to include several new minor contributors, such as China, Korea and Brazil.

Others point out that ODP's new Long Range Plan, which is under discussion, goes some way to addressing French and British concerns. Under the new plan, for example, the ship would spend longer periods at single sites than previously.

ODP is putting a brave face on the possibility that Britain and France might withdraw. "The French have rattled their sabre before", says Timothy Francis, the deputy director of ODP, arguing that the high costs of ocean drilling means that no one country or geographical bloc could realistically consider doing it alone.

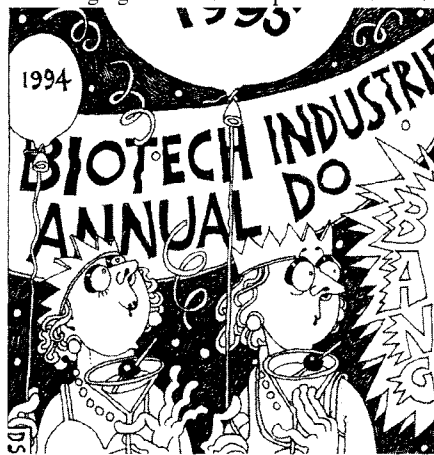
Declan Butler

A good year for US biotech, but caution is still advised

San Francisco. After the disappointing performance of biotechnology shares in the first half of last year, the US stock market has been cheered by an impressive recovery in the second half, contributing to an overall gain in stock value of 68.4 per cent over the whole year. But there remains a nagging sense inside and outside the industry that the good times may not last.

"If this is sustained, it'll be good for the industry," Costa Sevastopoulos, chairman of Metra Biosystems in Mountain View, California, said in San Francisco last week at the annual healthcare meeting organized by the investment company Hambrecht & Quist. But he warned that disappointment may be looming: too many areas of research, such as gene therapy and genomics, have overstepped their near-term potential value.

A depressed mood in the industry lifted last year in mid-June, when, after a string of poor results from clinical trials, the Philadelphia-based company Cephalon reported encouraging data from a phase II trial of



Myotrophin, its treatment for amyotrophic lateral sclerosis (ALS), better known as Lou Gehrig disease.

A rash of positive developments followed. Biochem Pharma of Montreal, began selling 3TC for AIDS; Gilead of Foster City in California filed for approval to market a broad-spectrum antiviral drug; and US Bioscience of Pennsylvania prepared to launch its treatment for advanced ovarian cancer patients.

Industry fortunes subsequently climbed. By the end of 1995, biotechnology companies had raised nearly \$2.2 billion from the public markets, compared to a little over \$1 billion in the previous 12 months. New collaborations with pharmaceutical companies increased in value to \$3 billion from about \$1.9 billion in 1994.

According to Dennis Purcell, managing director for life sciences at Hambrecht & Quist, advances in the industry have resulted

from the efforts of managers who learned some hard lessons the year earlier. Both the selection of drug targets and the design of clinical trials has become more careful, he said, while companies have cut overhead costs and learned to collaborate.

In all, more than 150 biotechnology products are undergoing phase III clinical trials, according to the investment company. "We're really taking the industry to a new stage," said Purcell. He added that the good news is likely to bring more. Many companies are flush with cash after negotiating link-ups with large drug developers, making them more attractive to investors.

In private, however, many industry leaders have doubts about the solidity of the basis for the market's enthusiasm. They warn that a new crop of young stock analysts and investors from other sectors may lack a sufficiently critical eye. Some of the drugs under development may not bring a significant improvement over existing treatments, they say, or may be too expensive for a cost-capped healthcare environment.

A recent merger between two biotechnology companies in south San Francisco, Arris Pharmaceutical Corporation and Khepri Pharmaceuticals, which broadened Arris' technology base and gave Khepri a way of gaining critical mass without going public, may provide a model for mid-size companies that want a hedge against an uncertain future.

Observers of the US biotechnology industry say that up to 40 similar companies could benefit from such relationships with others in their therapeutic area. Diminishing egos, resulting from the recent financial losses, and more financially demanding executive boards, are spurring on the consolidation.

Sevastopoulos sees a new pragmatism in the industry, with seasoned investors and quick-return venture capitalists replacing the early idealists. As a result, start-up companies are finding it harder to attract money, and new ideas are staying longer at the university level, or finding their way into existing companies.

But the second-generation managers, seasoned in the industry and experienced at making deals with major pharmaceutical companies, may contribute to improved sustainability, according to Sevastopoulos. "The arrogance of the past has dissipated," he says.

Sally Lehrman

INRA — Bernard Chevassus. The director of INRA is Bernard Chevassus, and not Roger Cassini, who recently resigned as the French representative to the European Commission's programme of agricultural research, FAIR (see *Nature* 378, 328; 1995).

MPs challenge rejection of genetics panel

London. A sharp conflict of views has arisen between the British government and a key group of Members of Parliament over the extent to which the government should take responsibility for mitigating the potentially harmful effects of new discoveries in human genetics and their social applications.

Last week, the government announced details of a new Advisory Committee on Genetic Testing — thought to be the first committee of its type in the world — which will have responsibility to ensure that genetic screening tests “are supplied safely and used ethically”.

At the same time, however, it rejected as “unnecessary” a proposal from the House of Commons’ Science and Technology Committee, made in a report issued last July after eight months of hearings, to establish a national Human Genetics Commission to maintain a strategic overview of the development of the science and the social and ethical issues it is likely to raise in the future (see *Nature* 376, 202; 1995).

In its response to the report, the government has said that it feels the existing network of “more targeted advisory bodies” is already “largely effective” in achieving this task. But members of the committee disagree and, in a highly unusual move, are now planning to hold a further hearing at which those previously called to give evidence will be asked to comment on the government’s response.

“Our concern is that without an overview of how genetics is developing, we could be faced with a series of *ad hoc* decisions engineered by successive crises,” says Spencer Batiste (Conservative, Elmet). “We need a flexible framework that can take a serious and measured overview of what is going on; by rejecting that proposition, the government has made a number of us on the committee unhappy.”

The government’s main reaction to the select committee’s report came from Ian Taylor, the science and technology minister, in answer to a parliamentary question. Taylor said that he welcomed the report’s support for the Human Genome Project, stressing that genetics provided “major opportunities for partnership between industry, the science base and government”.

Taylor also said that the government recognized “the sensitivities which can apply to the application of genetics in healthcare and other fields”. But the government’s written response indicates that it is reluctant to take extra steps to regulate such activities — for example, specifically rejecting the proposal that it should introduce legislation covering the use of genetic information by the insurance industry.

Unsurprisingly, this response met with relief within the industry. “I think that it is a sensible and balanced view by the govern-

ment,” says Paul Smea, of the Association of British Insurers, which had earlier been given one year by the science committee to draw up a code of conduct.

But the government’s response has come under fire from various groups which have been campaigning for more action to prevent the misuse of genetics in fields such

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

James Holmes/Collmark/SPL

Genetic identity: parliamentarians want a commission to monitor ethical issues.

as insurance and employment, and feel that government officials, while listening to their case, have ignored their conclusions.

“I am fairly dismayed,” says David Shapiro, executive secretary of the Nuffield Council on Bioethics, which last year produced a report on genetic screening referred to approvingly in the government’s response. “Our minimum recommendation was that the implications of what was happening in the fields of insurance and employment should come within the remit of an advisory committee. It is open to government to reject that view, but not to wave around the fact that they have taken our report into consideration in drawing their conclusions.”

The Genetics Interest Group, a pressure group which has been campaigning vigorously to defend the interests of those it considers vulnerable to ‘genetic discrimination’,

was even more outspoken in its criticism. Members of the group describe the report as a “missed opportunity” that failed to recognize the importance of transdepartmental links on genetics-related issues.

Kay Davies, professor of genetics at the University of Oxford, and chair of a committee that produced a report for the Office of Science and Technology two years ago on human genome research, arguing the need for a broad, strategic approach embracing both the interests of the public and the private sector, expresses similar concern. “If we do not get the infrastructure [for developing the applications of genetics] right now, the situation is going to become dreadful once we get on the issues raised, for example, by the ability to screen for a predisposition to multifactorial diseases,” says Davies.

For the present, however, the only new government activity will be to support the work of the advisory committee on genetic screening. Announced in parallel with Taylor’s statement by John Horam, the Parliamentary Secretary for Health, the committee will be chaired by the Reverend John Polkinghorne, president of Queen’s College, Cambridge, and a former physicist who headed an earlier inquiry into the use of fetal tissue in research.

The advisory committee will be responsible for monitoring the use — or potential use — of genetic screening tests within clinical practice, and as they are likely to be used through direct sale to the public. According to the Department of Health, the committee “will need to be reassured” on topics such as the implications of screening tests.

So far, government officials say they remain convinced that the various other issues raised by genetics can be adequately covered by similar existing advisory committees, as well as statutory bodies such as the Health and Safety Executive. But the members of the select committee are not expected to leave the rejection of their proposals unanswered. **David Dickson**

Charity head faces resignation calls

Paris. Jacques Crozemarie, chairman of France’s biggest charity, L’Association pour la Recherche sur le Cancer (ARC), is likely to be ousted at an emergency meeting next week of the charity’s executive board.

The meeting has been convened by the six-man working group set up earlier this month to study the claims of serious financial mismanagement at the charity made in a leaked report from the Cour des Comptes — the national audit commission (see *Nature* 379, 103; 1996). Their action was prompted in particular by Crozemarie’s sacking last week of Thierry Hercend, who had been managing director of ARC for just four months, apparently

on the grounds that he was being too cooperative with the working group.

Crozemarie himself announced during a radio interview last week that he would resign. Later the same day, however, he issued a statement saying that he would not resign until the full light was thrown on the “attacks” on ARC.

But the working group, which is made up of members of ARC’s executive board, has itself been accused of trying belatedly to distance itself from Crozemarie: the board has endorsed ARC’s practices in the past, and defended the charity against similar allegations to those made in the audit.

Declan Butler

Sodium leak blots Japan's nuclear prospects

Tokyo. The future direction of Japan's nuclear programme has been put in question both by the technical implications of last month's liquid sodium leak at the fast breeder reactor Monju, and by a growing scandal over an apparent attempt to suppress a video film of the accident scene.

The discovery that the leak was due to a fracture of one of many thermal probes in the sodium cooling system means that the reactor, at Tsuruga on the coast of the Japan Sea, may have to be closed down for several years for repairs and redesign.

In addition, moves by the Power Reactor and Nuclear Fuel Development Corporation (PNC) to prevent some revealing video footage taken shortly after the accident from reaching the media have undermined public trust in the corporation, which is responsible for implementing all research and development in relation to nuclear power in Japan.

Last week, an X-ray inspection of the leak site by a team from the Science and Technology Agency (STA), the government agency that administers PNC and Monju, revealed that liquid sodium in the secondary cooling system had leaked out through the fracture of a probe that measures the temperature of the sodium coolant.

The probe, which is encased in a stainless steel cylinder and extends about 18 cm into the liquid metal, apparently snapped off at a point where the cylinder rapidly tapers from 2 to 1 cm in diameter (see diagram). The liquid sodium leaked out through the broken cylinder. The STA experts do not yet know why the cylinder broke. It may have been the result of an isolated defect, such as a hairline fracture, in that particular probe, or, alternatively, to a more fundamental flaw in the design of the probes, of which there are 48 in the reactor's cooling system.

One theory under consideration is that vortices shed from the probe as the flow of liquid sodium varies between zero and 5 metres per second may have caused the probe to vibrate and eventually to fracture.

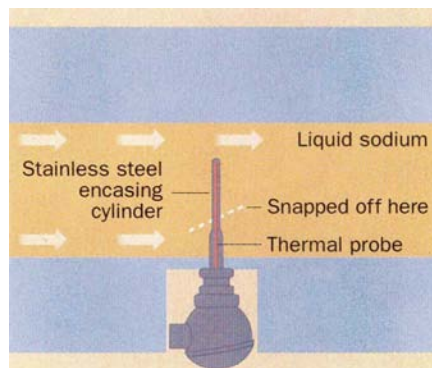
Similar probes were used for 20 years without problem in Monju's predecessor, the experimental fast breeder Joyo. But the Joyo probes were shorter and tapered more gradually, and PNC officials admit that no long-term vibration tests were carried out on the new design of probes used in Monju.

As serious as this admission was, an even greater public uproar has been caused by the revelation that PNC officials tried to hide from the public a videotape that they recorded a few hours after the accident in the room where sodium leaked, and that they also heavily edited other video footage that was released to the press and television.

Under pressure from local government officials, who videotaped the area of leakage for themselves three days later, PNC officials first admitted on 20 December that

they had edited a 4-minute video down to 1 minute before releasing it to the media, in order to remove sections showing sodium compounds dripping from the area around the thermal probe. This was apparently done in an attempt to play down the seriousness of the accident.

A subsequent investigation by the STA revealed that PNC had made another earlier video at about 2 a.m. on 9 December — a few hours after the accident — showing



A fracture of the temperature probe seems to have allowed liquid sodium to leak out.

officials wearing silver protective suits in a room filled with white smoke produced by the leaking sodium.

PNC officials had attempted to hide this video footage from the public. They had also apparently lied in a report to the STA stating that no-one had entered the room where sodium leaked until 10 a.m. that morning. Following this revelation, four local PNC officials at the Monju site were demoted and moved to new positions elsewhere in the corporation.

But the matter has not rested there. The results of an internal investigation by PNC released last Friday (12 January) has revealed that, despite claims by officials at the corporation's headquarters in Tokyo that they were unaware of the videotape taken by their local officials, a copy of the tape had in fact been viewed by ten head-

quarters staff the day after the accident.

After this latest admission, the STA decided to dismiss the head of PNC, Hiroshi Oishi. The next day, the PNC official in charge of the internal investigation, Shigeo Nishimura, leapt to his death from a hotel roof in Tokyo. Nishimura, deputy general manager of PNC's general affairs department, who was appointed by Oishi to head the investigation, is said to have written in a suicide note to Oishi, explaining that he was "exhausted by his duties". (It is a peculiarity of Japan that junior officials sometimes commit suicide if they play a role in getting their bosses into trouble, or if they feel burdened by knowing too much.)

It now seems likely, for both technical and political reasons, that Monju will not resume operations for at least several years. In particular, local government officials of Fukui Prefecture, where Monju is located, have emphasized that they will not allow operation of the reactor to be resumed until both they and the public are completely satisfied that the plant is safe.

The acting head of the Atomic Energy Commission, Yoshinori Ihara, has admitted that nuclear energy in Japan has "sustained a deep scar" with the Monju accident because it "lost the people's trust" in nuclear energy, in particular fast breeder reactors.

Japan has already had to delay and scale back its plans for recycling nuclear fuel and producing plutonium because of delays in the operation of Monju (see *Nature* 369, 596; 1994). Even so, the government is projecting that it will have about 70 tonnes of plutonium on its hands by 2010, and it plans to burn up much of this in conventional reactors by mixing it with uranium.

But the Monju accident has increased public resistance to the use of plutonium in any kind of reactor. Although the government has yet to take an official stance, it now seems inevitable that Japan will, at the very least, have to scale back and/or delay its plans for nuclear fuel recycling even further.

David Swinbanks

Russia appoints vice-premier for science

Moscow. Russian science has received some unexpected support from the Kremlin in the form of a decree from President Boris Yeltsin announcing the appointment of 50-year-old Vladimir Kinelev, head of the State Committee for Education, as a vice-premier for science and education.

It remains unclear what relationship Kinelev will have in his new post with Boris Saltykov, the Minister of Science and Technological Policy. Saltykov had been the vice-premier responsible for both science and education before the administration of these two activities was divided between the State Committee and the ministry.

Kinelev says that he was informed of his promotion by Viktor Chernomyrdin, the prime minister, rather than by Yeltsin himself. "An integrated system of science and education is being developed in Russia," Kinelev said last week. "Our target is to adapt this unique system quickly to the market economy, in order to create mechanisms for transferring new technologies to small and medium size business".

Science and education are expected to receive additional support in the near future from the recently elected State Duma, the lower chamber of the Russian parliament.

Carl Levitin

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

CERN photo

European debate on biotech highlights policy differences

Llewellyn Smith: CERN's director general led European delegation in Washington talks.

US role in LHC 'to be agreed within a year'

Washington. A mechanism for US participation in the construction of the proposed Large Hadron Collider (LHC) at CERN, the European Laboratory for Particle Physics in Geneva, Switzerland, will be agreed by the end of the year, according to negotiators at an initial meeting in Washington last week.

Officials from CERN led by its director general, Christopher Llewellyn Smith, and from the US Department of Energy and the US National Science Foundation agreed on the composition of a 'negotiating team' to oversee the detailed discussions of three working groups, which will start work almost immediately. They will deal with accelerator construction, scientific detector work, and the administrative and legal framework for an agreement respectively.

A communiqué issued by both sides said that CERN had stressed that it needed agreement on the scope of US involvement in the project's detectors by the end of 1996. CERN also said it needed to know the size of any US contribution to the construction of the accelerator by the end of 1997, when the laboratory's council is due to finalize the LHC's construction schedule.

In practice, the US budget process requires the Clinton administration to decide by the end of 1996 how much it would like to contribute, enabling Congress to consider the suggestion and, if it agrees, make the money available in the 1998 financial year, which begins in October 1997.

A US team will visit CERN in a few months' time to gather cost and schedule information on the project, and the full negotiating team will meet again in July to ensure the progress of the three working groups. The communiqué anticipates "that agreement will be reached within a year".

American support for the detector work is reasonably well-assured, but US involvement in construction will face opposition in Congress. The Department of Energy should, however, have some money available from 1998, when its domestic construction commitments start to decrease, and the construction work may win approval in Congress, provided the money is spent with US contractors.

Colin Macilwain

Brussels. The European Commission's new-found wish to be seen as a 'transparent' organization willing to mount open debates was manifested by a conference on biotechnology that it organized in Brussels last week. The meeting brought together in public for the first time representatives from Europe's three decision-making institutions — the commission, the European Parliament and the Council of Ministers — for an "exchange of views" in what is known in Euro-jargon as a *trialogue*.

The decision to hold the conference followed the unexpected rejection by the parliament last March of a draft directive from the commission seeking to clarify European law on the protection of biotechnological inventions, after the directive had already been approved by the council (see *Nature* 364, 103; 1995).

The directive would have confirmed that human genes and genetically altered animals and plants can be patented, and its rejection was the commission's first experience of the new powers granted to the parliament under the Maastricht Treaty. These introduced a process of 'co-decision' under which all three institutions, rather than just the council, which represents the governments of member states, must approve some European laws.

A revised version of the directive, adopted by the commission last month (see *Nature* 378, 756; 1995), is due to be formally published this week. The commission, anxious to improve its chances of being accepted by the parliament, organized last week's meeting "to promote an open dialogue on biotechnology between the institutions and more widely with all interested parties". The commission feels that closer collaboration with members of the European Parliament (MEPs), who have more direct contact with interest groups and pressure groups opposed to genetic engineering, could have avoided last March's defeat.

At the meeting, which was not intended to produce any common conclusions, Martin Bangemann, commissioner for industry, admitted that the commission can no longer conduct its business behind closed doors. "With the patent directive, we were unable to convince parliament," he said. "We can only do this if we are transparent."

Three further directives are under discussion in Europe in addition to that on patenting. A controversial draft directive on novel foods, which will have its second reading in parliament within the next few months, would require explicit labelling of foods produced by any new procedure — including biotechnological processes — that changes the chemical composition of the product.

The directive would not require labelling of foods produced by genetic engineering if they are chemically identical to those produced by conventional procedures. But many MEPs are calling for compulsory labelling of all genetically engineered foods.

The commission is also seeking to update two directives passed in 1990. Last month, it approved an amendment to the directive on the contained use of genetically modified microorganisms (GMOs), simplifying administrative procedures for those that are designated low-risk. This move is intended to reduce the burden of regulation on Europe's biotechnology industry, as is a parallel review of a companion directive on the release of GMOs into the environment.

The commission used last week's meeting to try to convince parliament that it is not ignoring important ethical issues in putting forward these directives. It introduced its Group of Advisers on the Ethical Implications of Biotechnology, a group of nine independent experts nominated by the commission, set up to advise the commission on ethical issues raised by biotechnology and to inform the general public of these issues.

But the group's work remains little known outside of the commission, and many MEPs argued strongly at the meeting that the commission had not in the past taken sufficient account of the ethical implications of applications of biotechnology, focusing instead on potential economic advantages.

Others expressed concern at the way that many MEPs stressed the potential risks, rather than the benefits, of biotechnology. Ian Taylor, for example, Britain's science minister, said that focusing on problems rather than benefits was wrong. "Biotechnology is important to Europe, and should not be hindered," he said.

In the past, the council of ministers has tended to be more enthusiastic about biotechnology than parliament. But there is also some dissent within the council itself. Denmark's Claus Grube, for example, told the meeting that his country was unhappy with both the proposal on labelling of novel foods and the draft amendment to the directive on the contained use of GMOs.

The diversity of opinions expressed at the meeting highlighted the difficulties that Europe is likely to face in the next few years in creating a revised framework for applications of biotechnological processes. There was a general welcome for the commission's efforts to bring the three-way debate into the open, but MEPs and council representatives alike complained about its poor organization. The agenda for the long-planned meeting was released only a few days before Christmas.

Alison Abbott

Livermore reopens plutonium building amid danger claims

San Francisco. The Lawrence Livermore Laboratory in California has reopened its plutonium building after a six-month shutdown following the discovery of major deficiencies in its safety practices. But the move has revived efforts by local critics to shut down the federal defence laboratory's activities involving plutonium machining and experimentation.

The laboratory announced in early December that it was calling workers back to the high-security plutonium building after the shutdown, which had led to between 30 and 40 workers being moved to other parts of the laboratory. Officials said they had implemented improvements designed to increase the safety with which the laboratory's 880 pounds (about 360 kg) of the radioactive material is handled and stored.

The facility was closed last April by the Defense Nuclear Facilities Safety Board after confirmation of safety deficiencies. These included the failure in early April to carry out an important inspection of the air pressure inside the building, and a lack of procedures to ensure the testing of other safety systems. The board also raised questions over safety training, inadequate staffing and ambiguities in safety procedures.

Laboratory officials say that they have investigated the problems revealed by the safety board, and have added improved safety equipment to the building, including a \$1.2-million emergency generator and a \$1.2-million ventilation system. Both the sprinkler and nuclear accident alarm systems have also been strengthened.

The workers who had been moved began returning to the plutonium area in October, and the facility reopened last month. But critics argued that there was no reason to reopen the building and to revive the hazard of plutonium handling on the site. "While the facility was shut down for six months, that was an opportunity to evaluate the need for it at all during the post-Cold-War era," says Marylia Kelley, spokeswoman for Tri-Valley Citizens Against a Radioactive Environment in Livermore.

Kelley points out that there have been continuing safety problems at the plutonium facility. She lists some of the potential dangers as earthquakes, fires, and the bulging storage canisters cited in a 1994 Department of Energy vulnerability assessment. She says the shutdown has given her group ammunition enabling it to press harder for a permanent closure, and that it has already submitted a petition with 3,000 signatures urging the halt of the plutonium manufacturing demonstration project. "We believe we are close to getting it shut down permanently," says Kelley. **Sally Lehrman**

Kidnappings 'must not deter' sponsors of research trips

London. Organizations backing a British student-led research expedition to Indonesia whose members were kidnapped last week by separatist rebels in the eastern province of Irian Jaya say the experience will not deter them from funding similar projects in the future.

Colin Bibby, conservation director of Birdlife International, which awarded £3,000 (US\$4,650) for the five-month project to catalogue some of the biodiversity of Irian Jaya, says his organization will review its procedures. But he says it would be a tragedy if fewer young people volunteered for overseas expeditions as a result of the experience of the expedition, known as Lorentz 95. "Most of our pioneering studies have been performed by young people," he says. "They are flexible, find it easier to overcome cultural barriers and can quickly get on with people."

The £20,000 expedition, which was organized by four recent graduates of the University of Cambridge, had also received advice and financial support from the Royal Botanic Gardens at Kew and the Royal Geographical Society (RGS).

Shane Winsor, head of the RGS's Expedition Advisory Centre, says she hopes the kidnappings do not put off parents from letting young people from taking part in research expeditions. Few of the 400 to 500 expeditions supported annually by the RGS are carried out by postgraduate researchers, she says. "In our experience, there is a greater risk from road accidents and malaria than from getting kidnapped."

Twenty-four expedition members — including local researchers, support staff and officials from the World Wide Fund for Nature (WWF) — were captured by the Free Papua Movement (OPM), a 30-year-old organization that has waged a low-key campaign for the region to break away from Indonesia and join with its geographical and cultural neighbour, Papua New Guinea.

The OPM is opposed to the Indonesian government's policy of transferring people from overcrowded parts of the archipelago to Irian Jaya. The movement also wants a larger share of the wealth from gold and copper mines on the southwest coast of Irian Jaya owned by US companies.

Irian Jaya, a Dutch colony until the early 1960s, is home to a number of native tribes with little interest in modern technology. They fight with wooden bows and arrows, and travel the waterways in canoes built from ancient designs. These people, however, are believed to be among the most linguistically sophisticated in the world.

Irian Jaya is one the few global regions

not to have been studied for its biological diversity — the last major expedition is believed to have taken place in 1911. Members of Lorentz 95 were working inside a 1.5-million-hectare region of Irian Jaya that is the subject of negotiations between WWF and the Indonesian government on designation as a national park.

But according to some observers, the OPM is concerned about plans for the park.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASON
REASONS

Cultural warriors? Tribes in Irian Jaya have been campaigning for independence from Indonesia.

Members of the movement fear that the Indonesian government may be using its tentative backing for a conservation area to consolidate its grip on the region.

The kidnappings have also raised the question of the safety of fieldworkers operating in sensitive parts of the world.

But Matthew Reynolds, a postgraduate student in development studies at the University of East Anglia and one of the founders of Lorentz 95, says the local political situation "was not considered to be dangerous." Reynolds, who was unable to travel to Indonesia for health reasons, says "Irian Jaya is half the size of Europe. The rebel groups are small, disorganized and equipped with bows and arrows. We figured the chance of coming across them would be very remote."

Bibby says the fact that Lorentz 95 was "an exceptionally well-thought out project" would not alone guarantee support. Backing from local authorities was essential, as was involvement from local researchers and the ability to work in safety.

Birdlife International, he points out, previously turned down two "first class research projects" to Rwanda and Burundi, well before the current civil wars, on the grounds that both regions were not considered safe to work in.

The organizers of Lorentz 95, he says, had good local contacts and made an effort to involve researchers from the Indonesian Institute of Biological Sciences. "They were better organized and better prepared than anything I've seen." **Ehsan Masood**

Witwatersrand seeks to defuse 'race row'

Cape Town. The University of the Witwatersrand last week agreed to lift a suspension it had imposed in December on its deputy vice-chancellor, William Makgoba, pending the results of a tribunal investigating allegations against him by 13 senior academics.

The suspension had followed Makgoba's decision to respond to the allegations — including the charge that he had misled the university over details of his previous career as a research immunologist (see *Nature* 378, 324; 1995) — by releasing details of the personal files of his accusers to the press, in particular hinting at possible tax evasion.

In agreeing to lift the suspension, the university appears to be hoping to take some of the heat out of what has become a highly charged controversy that has given rise to charges of racism against a liberal academic community.

Makgoba — tipped as a leading contender to become the next vice-chancellor — has undertaken to return to Robert Charlton, the current vice-chancellor, all copies or extracts from the personal files of the academics who signed the letter of complaint against him, on the understanding that he would have access to the information in the context of any future university

inquiry. He has also undertaken not to divulge confidential information outside the context of an inquiry.

Following a meeting last month with Charlton and Stephen Anderson, the chairman of the university's council, Sibusiso Bengu, South Africa's Minister of Education, had called for a "careful review" of Makgoba's suspension. The ruling African National Congress (ANC) had been less cautious, calling for Makgoba's immediate reinstatement and the use of mediation to address both sets of allegations.

At a meeting held on 12 December, the University of the Witwatersrand had responded by offering to lift Makgoba's suspension on conditions similar to those that he has now accepted, although they did not specify that he have access to the confidential information.

Makgoba initially rejected the university's offer, and two members of the council, Nthato Motlana, chairman of the company New Africa Investments, and Aggrey Klaaste, the editor of the newspaper *The Sowetan*, walked out of the meeting, because, it is said, of the way the issue was being handled.

It now appears more likely that Makgoba will appear before the three-member com-

mission. The third member of the tribunal will be Sir Colin Campbell, vice-chancellor of the UK University of Nottingham, who will join Lord Flowers and Walter Kamba, whose names were announced in late November.

Makgoba has said that he would only appear before a tribunal made up of individuals who were members of his own medical profession, and were his contemporaries. He demanded that they should also understand both his own background and the kind of transformation he claimed was needed at the university to meet the goals of post-apartheid South Africa.

But Makgoba subsequently rejected an offer by the council to try to accommodate this demand by appointing additional members. He is now holding discussions with a member of the council, Mr Justice Fikile Bam, and it was in an effort to facilitate these discussions that the council last week agreed to lift Makgoba's suspension.

In addition to the three-member tribunal, a separate commission, which is likely to be headed by Malcolm Wallis, chairman of the Senior Council of the Bar of South Africa, will investigate Makgoba's allegations against the 13 members of staff.

Michael Cherry

Congress steps back into dispute over Arizona telescope

Washington. An Arizona congressman has come to the rescue of the embattled Mt Graham International Observatory in Arizona, proposing an amendment to a congressional budget bill that would allow construction to proceed on the Large Binocular Telescope (LBT), the focus of a bitter controversy that has pitted astronomers against environmentalists and Native American tribes.

But action on the 1996 appropriations bill for the Department of Interior is still being held up, as other amendments have been rejected by President Bill Clinton. And opponents of the telescope, which is being built by a team of partners headed by the University of Arizona, warn that passage of the Mt Graham "rider" would further inflame the long-standing controversy over the telescope.

The tersely worded amendment, proposed by Jim Kolbe (Republican, Arizona), states that an alternative site for the telescope on Mt Graham currently being considered conforms to a law passed by Congress in 1988 allowing the university to build an observatory on the mountain.

Such a clarification would allow construction work, blocked by court order since May 1994, to resume without the university having to conduct a new environmental impact statement (EIS) for the alternative

site, which is 400 yards away from that which was originally proposed, as a local court has required it to do (see *Nature* 376, 719; 1995).

Opposition to the telescope originally focused primarily on its potential impact on the local population of the endangered red squirrel. But claims by the San Carlos Apache tribe that the mountain is sacred ground are rapidly becoming an even more contentious issue. Opponents of the telescope say that the mountain, called Dził Nchaa Si'An by the Apaches, should be excluded from further development under the Native American Religious Freedom Act — a matter that would have to be addressed in a second EIS.

The Kolbe amendment sidesteps the requirement for an additional EIS. But this has only helped to inflame critics such as Peter Warshall, a former biologist at the University of Arizona and a leading anti-telescope activist. Warshall says that if the amendment passes, "the anger level goes up tremendously, because the university [will have] avoided due process of law again".

Native American activists are also upset that Kolbe's rider was inserted as an alternative to an amendment to the Interior bill proposed by the Senate that would have required a study of ways to prevent HIV-AIDS trials to be

carried out on Indian populations.

Administration officials are expected to sit down with congressional Republicans in the near future in an attempt to hammer out their differences over the Interior appropriations bill. So far, however, the White House has not raised specific objection to the provision covering the construction of the Mt Graham telescope, and supporters expect it to remain in the final bill.

Peter Strittmatter, of the university's Steward Observatory, says that the proposed amendment would "give us the go-ahead" to complete the \$80-million LBT, which still needs another partner to provide \$15 million towards its construction costs (the university is courting potential partners in Germany and the United States).

But the controversy over the telescope has become as much a culture clash — 'big astronomy' versus eco-warriors — as an argument over squirrels and religious freedom, and as such is unlikely to end either quickly or easily. Thus even if Kolbe's amendment passes, more challenges are inevitable, says John Fitzgerald, a Washington-based lawyer who represents the anti-telescope Mt Graham Coalition: "We are not out of options."

Tony Reichhardt

Snow blocks US efforts to tackle research grant applications backlog

Washington. The main science-funding agencies in Washington have now been closed by political events or the weather for a month, throwing the federal grant application process into chaos. Snow closed the agencies for most of last week, and they expected to reopen for business only on Tuesday, 16 January, the day after the Martin Luther King holiday and 32 days since most of them were first shut down by political disagreement, on 15 December.

A break in the snow last Thursday gave officials at the National Institutes of Health (NIH), the National Science Foundation and the National Aeronautics and Space Administration a brief chance to survey the mountain of mail accrued since then. "It will be months before it will be business as usual," says Wendy Baldwin, deputy director for extramural research at NIH, which is postponing most planned grant review meetings up to mid-February, to allow time for adequate preparation. □

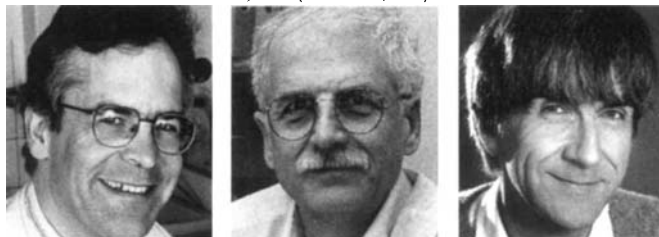
Japanese astronaut retrieves raft

Tokyo. Koichi Wakata, the Japanese astronaut on the US space shuttle Endeavour, successfully completed a major experiment for Japan's space agencies on 13 January when he plucked a re-usable orbital platform, the Space Flyer Unit (SFU), out of orbit with the shuttle's robot arm. The SFU was launched in March to carry out experiments including materials testing for the Japanese Experimental Module of the US space station, growth of crystals under microgravity, and breeding of newts under zero gravity (see *Nature* 374, 295; 1995).

At one point the SFU's solar panel failed to close properly as the shuttle closed to within tens of metres of the platform. The problem was overcome by jettisoning the panel, and Wakata was able to manoeuvre the \$0.4-billion platform into the cargo hold. □

Three awarded prize for medicine

London. The 1996 Louis-Jeantet prize for medicine will be awarded to Björn Dahlbäck (below, left) of the University of Lund, Sweden, Ulrich Laemmli (centre), of the University of Geneva, Switzerland, and Nigel Unwin (right) of the Medical Research Council Laboratory of Molecular Biology (LMB), United Kingdom. The winners will each receive SFr600,000 (US\$517,000) towards their research



and a personal award of SFr100,000. The prize — awarded annually since 1986, and named after its benefactor, a wealthy French businessman who died in 1981 — will be presented at a ceremony in Geneva in April.

Unwin has collaborated with a past prize-winner, his colleague Richard Henderson, director designate of the LMB (see p193). □

France limits embryo implantations

Paris. The Paris Supreme Court of Appeals last week rejected an appeal by a woman seeking to be implanted with an embryo created by her and her husband before the latter's death in an accident last year. The court's ruling in the test case confirms that *in vitro* fertilization in France is limited to "living, sterile couples (a male and a female) of reproductive age".

The woman has described the situation as a "revolting" infringement of personal liberty and questioned the ethics of the court's ruling that the embryos could be given to another couple, pointing out

I'm looking to purify peptides,
proteins and oligos simply,
accurately and fast.

Whether you're purifying a peptide, protein or oligonucleotide, one tool assures you of receiving your results simply. The tool is pre-packed RESOURCE™ columns from Swedish-based Pharmacia Biotech. You'll find every RESOURCE leads to efficient and convenient working methods for a wide range of applications.

Efficiency is yours as RESOURCE allows for excellent resolution at high flow rates with high capacity at moderate back pressures. And convenience is yours as you can work with ion exchange, RPC or HIC.

What gives RESOURCE its combination of simplicity, accuracy and speed? The answer is the media. Every RESOURCE is pre-packed with SOURCE™ media—so while pre-packed RESOURCE will save you the time of doing your own packing, monosized beads of SOURCE assure you of reproducible results and batch to batch consistency.

With this blending of media with uncompromising pore structure and a tool with convenient handling and working procedures, pre-packed RESOURCE offers a unified approach for complete small scale purification. To find out more about RESOURCE, just call us at 1 (800) 526 3593 in the United States or +46 18 16 5011 from the rest of the world. It's that simple.



that this could create a situation where she might one day recognize her children as part of another family. But the court's decision has also emphasized the lack of consensus on bioethical issues within Europe. If the woman had lived in the United Kingdom, implantation would have been allowed, if approved in her husband's will, according to the UK Human Fertilization and Embryo Authority. □

Russian inquiry into abortion claim

Moscow. The Russian prime minister, Viktor Chernomyrdin, has ordered the minister of health Aleksandr Tsaregorodtsev to investigate claims made in a television programme that Russian doctors have been involved in aborting late-stage fetuses in order to manufacture medicines used to treat children suffering from disorders such as Down's syndrome.

According to the news agency ITAR-TASS, the inquiry follows allegations made in the television programme *Itogi*, broadcast on 7 January, that doctors were involved in the preparation of a chemical formula artificially to abort late-term pregnancies. □

Germany funds compounds library

Munich. The German federal research ministry has agreed to provide DM1.5 million (US\$1.04 million) to the Hans-Knöll Institute for Natural Product Research at Jena, in east Germany, to establish a library of compounds isolated from natural sources. Five German pharmaceutical companies will invest an additional DM1.5 million in the venture, in return for the right to screen the products for potential pharmacological activity. □

Greens oppose enriched uranium

Munich. The environmentalist group Greenpeace International last week joined the campaign to prevent highly enriched uranium (HEU) being used in a research reactor planned by the Technical University of Munich in Germany. A report from the group warns that a contract to supply Russian HEU for the reactor,

known as FRM2, could be signed by the end of the month.

According to the report, secret negotiations between the European atomic energy agency, Euratom, and its Russian counterpart, Minatom, have been taking place since September 1995. Greenpeace says that any agreement with Euratom could make it easier for Russia to sell HEU to other countries. □

Europe to share research ships

Paris. France, Germany and the United Kingdom will hold a tripartite meeting in Bonn next month at which they are expected to sign an agreement to share the use of each other's research vessels.

According to an official from the UK Natural Environment Research Council, it is hoped that the agreement will lead to the creation of a single European research fleet within the next 15 years. Europe currently has 40 research ships — half of which are either British or French — and two research submarines, both of which are French. □

Sex determination banned in India

New Delhi. The use of sex determination techniques such as amniocentesis to identify female fetuses for abortion became illegal in India on 1 January. Informing pregnant mothers or their relatives of the sex of a fetus is also specifically prohibited by the law, which was passed by parliament eighteen months ago. The law allows prenatal tests to be carried out only in registered laboratories "for the detection of chromosomal abnormalities and genetic disorders and congenital abnormalities". The use of this knowledge for selective female feticide will be punished with a fine of Rs10,000 (US\$300) and up to three years in prison.

Any doctor who carries out illegal tests is also liable for punishment. According to the Indian government, the declining sex ratio (927 females per 1,000 males, as recorded by the 1991 census) "has been a matter of serious concern". The new law has been introduced to "tackle the causes of gender imbalance". □

Here's a handful
of simple RESOURCES
from Sweden.



**Pharmacia
Biotech**

Uppsala, Sweden. (And the rest of the world)

Crystallographic data deposition

SIR — A formal discussion of the archival journal requirements for data deposition was held at the international seminar-cum-school on macromolecular crystallographic data at Calcutta, India, in November.

The current policy of the International Union of Crystallography (IUCr) is that on publication of a crystal structure determination of a macromolecule, the atomic parameters used or represented in the publication must be deposited in the Protein Data Bank. The deposition of structure amplitudes is recommended but not insisted on. The policy provides crystallographers with the option to delay the release of atomic parameters for one year and of structure amplitudes for up to four years from the date of publication. Participants strongly supported this policy and felt it should be strictly applied by the journals.

But recent developments in X-ray crystallographic experimental and refinement techniques and the huge expansion in computer power and networking make it necessary to review deposition arrangements. It was noted that the new validation procedures are much more effective but require the experimental structure amplitudes as well as the atomic parameters. In addition, the technical arrangements for deposition, analysis and validation of macromolecular crystal structures are now much easier.

We consider it vital for the macromolecular crystallographers to respond to these developments in their deposition practices. We recommend, therefore, that publication of macromolecular crystal structures should be accompanied by deposition of atomic parameters and also structure amplitudes. Among the many reasons identified for this practice, the following are critical.

(1) Rigorous validation of the structure determination results can be carried out only by using both atomic parameters and experimental structure amplitudes. It is important that journals should ensure that referees have sufficient information to prevent incorrect structures from being published.

(2) Archiving of these data will ensure that they are not lost. Numerous reports at this meeting of data being lost probably reflect a general problem in the crystallographic community.

Edward N. Baker (Member of IUCr Executive Committee & IUCr Commission on Biological Molecules), Department of Chemistry and Biochemistry, Massey University, Palmerston North, New Zealand

Tom L. Blundell, ICRF Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, London WC1E 7HX, UK

Mamannamana Vijayan (Chairman of IUCr Commission on Biological Molecules), Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Eleanor Dodson (Member of IUCr Electronic Publishing Committee), Department of Chemistry, University of York, York YO1 5DD, UK

Guy Dodson (Previous Chairman of IUCr Commission on Biological Molecules), Department of Chemistry, University of York, York YO1 5DD, UK

Gary I. Gilliland, Center for Advanced Research in Biotechnology, Rockville, Maryland 20850, USA

Joel L. Sussman (Head, Protein Data Bank), Departments of Biology and Chemistry, Brookhaven National Laboratory, Upton, New York 11973, USA & Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel

■ See also Opinion, page 191.

Panic and the Pill

SIR — The Committee on the Safety of Medicines was indeed surprised by some aspects of the media's response to its advice on oral contraceptives (OCs) containing gestodene and desogestrel (*Nature* 377, 663; 1995). Our particular concern was that the lay press might publish misleading articles leading to a 'pill scare' with many women stopping using oral contraceptives altogether. The consequence would have been unplanned pregnancies, an increase in the abortion rate and the various risks that these entail. In fact, the lay media did an excellent job of reporting accurately the relevant brands, and of passing on our message that no one need stop taking OCs. In contrast to our expectations, it was the scientific press that failed to get the facts right.

Having correctly emphasized the importance of absolute risk, you quote a figure (1 in 200,000 per year) that is at least 30 times less than any estimate ever published for thromboembolism with OCs. The best estimate we have is that the excess risk associated with using OCs containing desogestrel or gestodene is about 15 thromboembolic events per 100,000 per year. Although the lay media quoted the correct figures for absolute risk, *Nature* was not alone in providing its readers with incorrect factual information. J. Guilleband in the *British Medical Journal* (311, 1111–1112; 1995) managed to express the excess risk as 15 per 100,000 women. Because the risk remains fairly constant over time, and because women tend to use OCs for many years, Guilleband also seriously underestimated the absolute risk. Once your own estimate of risk is multiplied by 30, the whole thrust of your leading article, that the Committee on the Safety of Medicines (CSM) acted unnecessarily, becomes untenable.

There are other serious inaccuracies. The implication that the committee and the

Medicines Control Agency failed to communicate with the relevant pharmaceutical companies is false; and the suggestion that they might be best placed to put out warnings that have important implications for public health is naive. It is also wrong to suggest that Professor Walter Spitzer might be in a better position to judge the position than the committee. That could be true only if he had information that he had not shared with us — and that is not the case. The CSM's advice was based on all the information available from three separate studies, and Spitzer's data were only part of the picture. Our advice also took into account considerations relating to contraceptive use and practice in the United Kingdom. Needless to say, I was surprised by Spitzer's actions; he presumably made the data available to us because he felt they might have important implications for public health. It is therefore difficult to understand why he objects when they were used for the purposes of protecting public health.

Finally, I find your statement that "sudden announcements... are almost always likely to cause panic" quite extraordinary. The press notices emphasized that there was no need to panic, and this was clearly explained on British television. How on Earth can an announcement like this be made gradually?

Michael D. Rawlins

(Chairman)

Committee on the Safety of Medicines, Market Towers, 1 Nine Elms Lane, London SW8 5NQ, UK

Games people play

SIR — Sally Lehrman errs in saying that the University of California at San Francisco (UCSF) rivals Stanford University "on the football field" (*Nature* 378, 529; 1995). UCSF has never had a football team. Stanford cherishes its football rivalry with the University of California, Berkeley, known also as 'CAL' and 'The Golden Bears', the latter in reference to the California Grizzly, totem animal of CAL and the State of California. Stanford shows its preoccupation by a toll-free telephone number 1-800-BEAT CAL. The corresponding UC number is 1-800-GO BEARS. The University of California, Los Angeles, also has a football team that frequently trounces Stanford.

Thomas H. Jukes

University of California, Berkeley, Berkeley, California 94720, USA

Correspondence

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to nature@nature.com

The wake of a legendary earthquake

Hiroo Kanamori and Thomas H. Heaton

Although quiet this century, the Cascadia subduction zone has produced big earthquakes in the past. But how big? An ingenious study involving documents written 300 years ago suggests that they can be giants.

WHETHER a giant earthquake could occur in the Cascadia subduction zone, off the Oregon and Washington coast of North America, has been a subject of hot debate for the past decade because of the tremendous destructive potential of such an event. Many lines of evidence, from plate tectonics to sea-floor sediments, have been used to argue 'convincingly' both for and against the possibility of giant earthquakes in the region. Evidence of extensive coastal subsidence of at least 1 metre, dated by carbon-14 and by tree-ring studies, suggests that the last such earthquake may have occurred about 300 years ago¹. An uncertainty in timing of about 20 years seemed as good as one could get for these old events. However, a study by Satake and colleagues² (reported on page 246 of this issue) combines an investigation of old Japanese documents with computer simulations of tsunamis to determine the size of this event and its timing — not only the date, but even the hour. As Satake *et al.* mention, this timing seems to agree with American Indian legends saying that coastal flooding occurred on a cold winter's night.

In the past century no subduction zone in the world has exhibited such low seismic activity as Cascadia³ — no earthquake with moment magnitude (M_w) larger than 6 is known to have occurred there for the past 70 years. It is human nature not to worry too much about seismic hazard if the present level of activity is very low. However, between the late 1960s and the early 1980s, as the study of plate tectonics and giant earthquakes ($M_w > 9$) advanced, Cascadia began to look less safe. The oceanic plate there is young (about 10 million years old), and large amounts of sediment are dumped on the trench by the Columbia river. Other sites of giant earthquakes have similar features, suggesting that giant earthquakes could also occur in the Cascadia subduction zone⁴.

In opposition to this view some scientists argued that Cascadia is different because of the extreme youth of its subducting plate, and that the combination of probable high temperatures and the quantity of water-saturated sediment deposited there would effectively lubricate the plate, and prevent the level of strain accumulation that produces giant earthquakes. This controversy excited the interest of geologists, and extensive

efforts began in the late 1980s to identify geological evidence for giant earthquakes in the area⁵. They found alternating layers of peat and mud in estuarine sediments at several places along the Cascadia zone, a pattern that indicates sudden subsidence and submergence by the ocean. The subsidence could be interpreted as due to crustal deformation during a giant earthquake. Similar subsi-

with $M_w \sim 7.5$ would cause very strong ground shaking, but its effect would be limited to a distance of about 200 km from the rupture zone. Tsunami height from such earthquakes rarely exceeds 2 metres. In contrast, an $M_w = 9$ earthquake could rupture the entire 1,000 km length of the subduction zone, and would cause strong ground motion up to a distance of several hundred kilometres from the rupture

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Seattle, in Washington state. The city is about 150 km from the Cascadia rupture zone. Anchorage, at roughly the same distance from the Alaskan subduction zone, was strongly shaken by a magnitude 9.2 earthquake in 1964.

dence occurred during the two largest earthquakes this century — in Chile in 1960 ($M_w = 9.5$), and in Alaska in 1964 ($M_w = 9.2$). Each subsided peat layer is overlain by a distinct layer of sands which look like those deposited by tsunamis; similar features were found after the 1960 Chilean earthquake. The palaeoseismic studies were followed by tree-ring studies, which document the widespread death of trees near sea level about 300 years ago^{6,7}.

These observations have convinced most scientists that during the past 2,000 years the region has experienced at least several large earthquakes. However, one of the important questions was whether these earthquakes were really giant earthquakes with $M_w > 9$, or a series of somewhat smaller earthquakes with $M_w > 7.5$ (ref. 8). In terms of the implications for seismic hazard, especially tsunami hazard, the distinction between these two possibilities is very important. An earthquake

zone. The maximum tsunami height would probably exceed 20 m, and considering the large population in the coastal areas of Oregon and Washington, such huge tsunamis would have devastating effects.

With all the limitations of comparative and palaeoseismological studies, it seemed that reducing these uncertainties would be extremely difficult. The study of Satake *et al.* has changed that view. In tsunami catalogues and old Japanese doc-

1. Atwater, B. J. *Geophys. Res.* **97**, 1901–1919 (1992).
2. Satake, K., Shimazaki, K., Tsuji, Y. & Ueda, K. *Nature* **379**, 246–249 (1996).
3. Heaton, T. H. & Hartzell, S. H. *Bull. seism. Soc. Am.* **76**, 675–708 (1986).
4. Heaton, T. H. & Kanamori, H. *Bull. seism. Soc. Am.* **74**, 933–941 (1984).
5. Atwater, B. F. *Science* **236**, 942–944 (1987).
6. Atwater, B. F., Stuiver, M. & Yamaguchi, D. K. *Nature* **353**, 156–158 (1991).
7. Nelson, A. R. *et al.* *Nature* **378**, 371–374 (1995).
8. Hyndman, R. D. & Wang, K. J. *Geophys. Res.* **100**, 22133–22154 (1995).
9. McCaffrey, R. & Goldfinger, C. *Science* **267**, 856–859 (1995).

Robert Harding

uments they found reports of flooding that occurred at several places in Japan on the night of 27/28 January 1700. There was no large earthquake at that time in Japan, Kamchatka, the Aleutian-Alaska region, or South America. By elimination, they concluded that these inundation events must be due to tsunamis caused by a giant earthquake in Cascadia. From their computer simulation of tsunami production and propagation, they found that it must have happened at about 21:00 on 26 January, local time, and had a magnitude of about 9.

Interpreting old documents is not as straightforward as it sounds, especially in this case. Japanese has changed considerably since the 17th century, and descriptions are often subtle and vague. Reading and interpreting them properly is a special skill, which Satake and colleagues have

been able to combine with the more modern skills required in making rigorous computer simulations. It is probably fair to say that some uncertainty remains in their interpretations, and further studies should be made of the behaviour of water-saturated deformable sediments and the complex deforming upper plate of the Cascadia subduction zone⁹. Nevertheless, the evidence presented by Satake *et al.* appears strong enough to justify serious studies of how best to prepare for a giant earthquake in Cascadia, and for the tsunami that would follow. □

Hiroo Kanamori is in the Seismological Laboratory, and Thomas H. Heaton is in the Division of Engineering and Applied Science and the Seismological Laboratory, California Institute of Technology, California 91125, USA.

LINGUISTICS

Making sense of rongorongo

Paul G. Bahn

It is not often that *The Journal of the Polynesian Society* and *The Rapa Nui Journal* feature in these pages. That they do so now is because of two papers by S. R. Fischer^{1,2} that help resolve one of the enduring enigmas of Easter Island, the tiny speck of land lost in the South Pacific. What is the meaning of its unique *rongorongo*, one of the world's few undeciphered scripts?

The script comprises parallel lines of characters, many of them bird symbols, hooks and so on (see figure), engraved on wooden tablets. Every other line is upside down, and the overall impression is of a tightly packed mass of uniform, skilfully inscribed hieroglyphics. Not one of the early European visitors who came to the island after its discovery by the outside world in 1722 ever mentioned the tablets, even though some spent days exploring ashore and entered native houses. The earliest written record of them is by a missionary in 1864, who said they were to be found in every house. Later visitors reported that they were kept apart in special houses and were strictly taboo.

It seems most likely that *rongorongo* was a very late phenomenon, directly inspired by the visit of the Spanish in 1770, when a written proclamation of annexation was offered to the chiefs and priests to be "signed in their native characters". This was probably their first experience of speech embodied in parallel lines, and they then adopted a method of

script that employed motifs they had already been using in their rich rock art³.

The script now survives only as markings on 25 pieces of wood, scattered around the world's museums, which contain over 14,000 'glyphs'. Since the Peruvian slave raids of 1862 removed the last aristocratic or priestly islanders who could truly understand the tablets, their content

sent an alphabet or even syllables, but were 'cues' for whole words or ideas, and a means of keeping count, like rosary beads. Since Barthel's work in the 1950s, *rongorongo* has largely been the preserve of the lunatic fringe, as well as of a series of bona fide scholars in Russia and elsewhere.

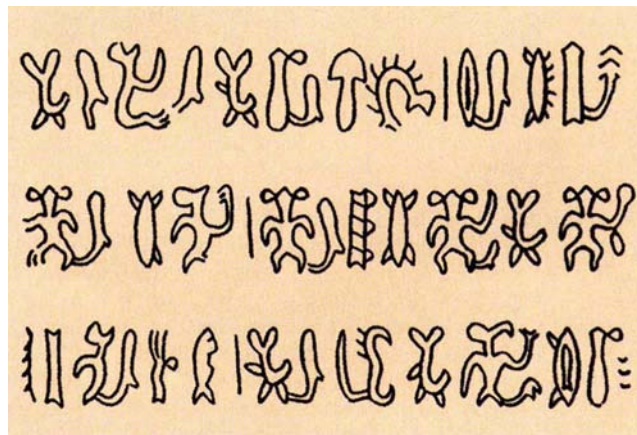
The new developments come from Steven Fischer, based in New Zealand, who has examined almost all of the tablets at first hand. His Rosetta stone was the Santiago staff¹, a 2-kg wooden sceptre acquired by Chileans in 1870; measuring 126 cm by 6.5 cm, it once belonged to an Easter Island *ariki* or leader, and is now housed in Santiago's Natural History Museum. Fischer discovered that, uniquely among *rongorongo* inscriptions, it marks textual divisions with about 97 irregularly spaced vertical lines (see figure). He then noticed that each glyph that starts a new division (that is, is immediately to the right of a vertical line) is suffixed with a phallus-like motif.

Fischer then observed that, in the series of glyphs within each division, almost every third one (the 4th, 7th, 10th, 13th and so on) also has this phallic suffix. Not one division has the suffix on its last or its penultimate glyph; not one division contains fewer than three glyphs; and most divisions are multiples of three, with the first in each trio sporting a phallus in almost every case. In other words, the Santiago staff has a basic triad structure.

It turned out that the same structure occurred — without the helpful vertical lines — on two other *rongorongo* artefacts, the Small Santiago tablet and the Honolulu tablet I. This led Fischer to characterize the structure by the formula X'YZn: X designates the glyph with phallus suffixed; Y is a statistically more frequent glyph that follows X; Z is a statistically less frequent glyph that follows Y; and n is the constant, denoting unspecified repetition of the triad structure.

In the light of this structure, Fischer has made a tentative reading, based on information provided by an islander in 1886, who sang a creation chant, '*Atua Mata Riri*, which was probably pre-missionary in origin, although linguistically contaminated, and almost certainly based on an original *rongorongo* composition. The chant — a list of 41 fanciful copulations and issues, such as "God Mata Riri copulated with Sweet Lime; there issued forth the poporo plant" — also has a triad structure, X'YZ. Here the copulator is X; the ' is the phrase "copulated with", equivalent to the phallic suffix; the copulatee is Y; and Z is the issue.

This is a common structure for ancient



Detail from the Santiago staff, showing a few of the elements of *rongorongo* script. (Adapted from ref. 1.)

has remained a mystery. For many years, the foremost expert on *rongorongo* (meaning 'chants' or 'recitation') was the German scholar Thomas Barthel⁴ who found about 120 basic elements in the glyphs (mostly stylized objects or creatures), combined to form between 1,500 and 2,000 compound signs. Barthel concluded that this was a rudimentary phonetic writing system, with picture symbols expressing ideas as well as objects — in other words, the individual glyphs did not repre-

Polynesian creation chants and genealogies: male copulates with female and produces offspring. Fischer therefore concludes that the three *rongorongo* inscriptions on which he has detected the triad structure are cosmogonies (creation chants) — not genealogies, since in these the issue (Z) would have to be the subject of each subsequent triad, which only occurs three times on the Santiago staff, and not at all on the others.

In short, Fischer appears to have securely defined the genre of at least three of the *rongorongo* inscriptions. He also suggests that it is a mixed writing system, being both logographic and semasiographic — logographic because X represents a physical object, a word or words (such as '*Atua Mata Riri*, "God angry eyes", in the 1886 chant), while the phallic-like suffix is semasiographic, providing direct visual communication of the phrase "copulated with" — that is, it depicts an act without recourse to language. He provisionally deciphered one triad from the Santiago staff to read "All the birds copulated with fish, and there issued forth the sun".

In a second development², Fischer has found that most surviving *rongorongo* inscriptions — perhaps even 15 of the 25 — also consist wholly or in part of cosmogonies or procreation triads. He first noticed that the bird/fish/sun triad from the Santiago staff is repeated on a different tablet, known as Echancrée, except that the bird glyph has no phallus (this is seen as a graphic elision, signifying an evolution or variation during the script's short history). He therefore assumed that this tablet also comprised a sequence of procreations and hence a cosmogony. Similar cosmogonies were then found on 11 other tablets, all of them lacking the phallic suffix on the X glyph of each triad. Because these artefacts survived at random, it can be supposed that most *rongorongo* inscriptions now destroyed or lost also constituted procreation chants.

Fischer's conclusions also mean that during its period of use — probably between 1780 and 1865 — *rongorongo* was not merely a mnemonic device to aid in the recall of previously memorized texts, but had to be physically read. In other words, it was indeed a form of writing in which the priests composed creatively — the only known indigenous script in Oceania before the twentieth century. □

Paul G. Bahn, a freelance writer on archaeology, is at 428 Anlaby Road, Hull HU3 6QP, UK.

Martians in a deep freeze

Horton E. Newsom

HAS there ever been life on Mars? The latest consensus is that historical conditions on the planet as a whole were colder and even more inimical to water-based life than had been thought. Hopes of finding life had already been dented by negative results from the experiments on board the Viking landers, twenty years ago. Although examination of the martian rocks that have bombarded the Earth throughout its history^{1,2} has revealed evidence of carbonate deposition, presumably from aqueous fluids at or near the martian surface, and one of the rocks recently identified contains polycyclic aromatic hydrocarbons³, none of these meteorites has revealed evidence of biological activity. However, biological studies of terrestrial environments such as hydrothermal vents on the sea floor, ice-covered Antarctic lakes, deep aquifers, and the microbial communities beneath the surface of Antarctic rocks, have encouraged interest in the prospects of finding life in similar environments on Mars. The continuing search for life and the spectre of biological hazards from a martian sample return were discussed at several sessions of a meeting in San Francisco last month*.

Twenty-five years after images from Mariner 9 revealed dendritic valley networks whose formation was attributed to runoff from precipitation, a new view of the early history of Mars has emerged, with a colder early environment, but one that still had temporary rivers and lakes on the surface. The presence of water in underground aquifers was inferred from the large outflow channels, and from the lobate ejecta surrounding rampart craters. These observations led to speculations about an early, dense, carbon-dioxide-rich atmosphere that would produce a warm climate with an Earth-like hydrologic system, including rain, rivers, lakes and possibly even oceans. This view has given way to a colder, more Arctic picture, because such a greenhouse effect is unlikely to have fully offset the reduced heat flux from the early Sun⁴. In addition, careful studies find that the distribution and morphology of valley and channel systems are not consistent with the action of rain water. However, a minority of scientists

still favour brief warm, wet periods due to occasional orbital and obliquity changes, that may even have produced transient oceans.

Thus, studies of martian geology and inferred climate history suggest that before about 3.8 billion years ago the planet had a cool climate but an abundant supply of groundwater, melted by geothermal and impact-related heat sources. The dendritic valley networks probably formed by the dislodging of debris at

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The extinct volcano Olympus Mons, flanked by cloud. Heat from active volcanic sites on Mars might provide temporary, local environments with liquid water and suitable for life.

valley heads due to sapping, a weakening process that occurred at the site of springs supplied by large aquifers located under a frozen 'cryosphere'. Many of the valleys have central ridges and flat profiles, suggesting that material was largely transported down the valley not by water but by 'mass wasting', where material is moved by gravitational forces, possibly in the form of ice-cored rock glaciers (M. H. Carr, US Geological Survey, Menlo Park).

At the end of this early period, the rate of erosion dropped by a factor of a thousand, and Mars entered a long frozen era. Early basin-forming impacts on Mars would have destroyed any life forms near the surface. However, lakes favourable for the development of life could have formed in impact craters >10 km in diameter throughout the history of Mars, heated by the thermal energy in shocked and melted rock. Also, the heat from volcanic events could have created hydrothermal plumes that supplied water to the channels observed on the flanks of some martian volcanoes.

Over the next decade the revised picture of Mars will be tested by a flotilla of spacecraft. Launches planned before the end of 1998 include two US orbiters, two

1. Fischer, S. R. *J. Polynesian Soc.* **104**, 303–321 (1995).

2. Fischer, S. R. *Rapa Nui J.* **9**, 99–107 (December, 1995).

3. Bahn, P. G. & Flenley, J. R. *Easter Island, Earth Island* 203–207 (Thames & Hudson, London/New York, 1992).

4. Barthel, T. *Grundlagen zur Entzifferung der Osterinschrift* (Cram, De Gruyter, Hamburg, 1958).

*American Geophysical Union Fall Meeting, San Francisco, USA, 11–15 December 1995.

US landers, and one Russian and one Japanese orbiter. These spacecraft will also search for signs of hydrothermal activity related to volcanoes or impact craters. Such environments may be favourable for preserving evidence of microorganisms or prebiotic organic chemistry, in fine-grained stable minerals such as silica, phosphate, carbonate, and metal sulphides. The identification of such targets for direct study is central to the exobiology strategy⁵, and the search for microfossils at these sites offers the best hope of finding complex organisms (J. D. Farmer, NASA Ames Research Center).

On the horizon is the prospect of returning samples of martian atmosphere, soil and rock to the Earth. The new technology of 'in situ' resource utilization' will enable a small Delta-class rocket to launch a relatively inexpensive sample-return mission, which will produce from the martian atmosphere some of its own propellant for the return trip, possibly as soon as 2005. But what of the possible hazards associated with sample return (D. L. DeVincenzi, NASA Ames Research Center)? The contamination of the Earth with living martian organisms is undoubtedly unlikely, and of course there is the

argument that they would be as vulnerable to terrestrial life as the martians in H. G. Wells's *War of the Worlds*. Nonetheless, the possible legal barriers to a sample return that could be raised by citing environmental hazards, and the numerous federal agencies that could claim jurisdiction, are important issues. Given the possible existence of life in previously unconsidered environments, the possibility of life detection will never go away. The chances of finding evidence for past life on Mars are extremely remote, but the implications of making such a discovery ensure that the search will remain one of the most exciting aspects of the exploration of Mars. □

Horton E. Newsom is at the Institute of Meteoritics, Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA.

1. Melosh, H. J. *Planet. Rep.* XIV No. 4, July/August, 16–19 (1994).
2. McSween, H. Y. Jr *Meteoritics* **29**, 757–779 (1994).
3. Thomas, K. L. et al. *Lunar and planet. Sci.* **XXVI**, 1409–1410 (1995).
4. Squyres, S. W. & Kasting, J. F. *Science* **265**, 744–749 (1994).
5. Carr, M. H. et al. *An Exobiological Strategy for Mars Exploration*, NASA SP-350 (1995).

NEUTRON STARS

A faint light in the graveyard

John C. L. Wang

STARS too that have suffered violent deaths may come back in ghostly form. The solid remains are neutron stars, superdense objects a few kilometres across that are created when massive stars end their lives in spectacular supernova explosions. Some have stellar partners, whose substance they consume to become bright accreting binaries; others are isolated. Over the lifetime of the Milky Way, the Galactic graveyard is supposed to have been populated by huge numbers ($\sim 10^8$ to 10^9) of these isolated neutron stars. Whereas the new members of this graveyard reveal their existence quite

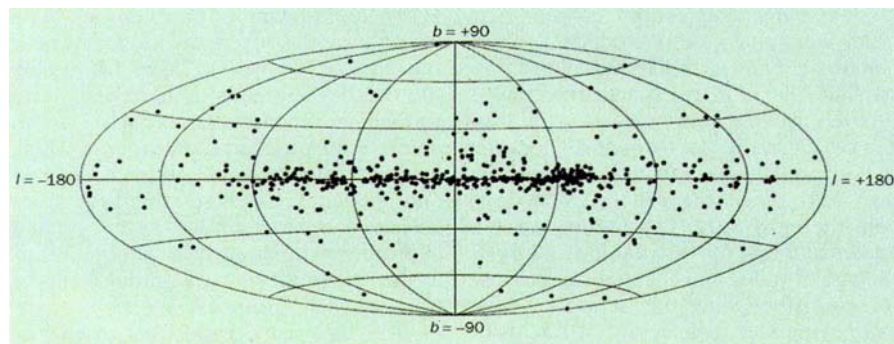
readily as radio pulsars (see figure), the older members ($\geq 10^8$ years old), which should be far more numerous, have remained eerily silent — that is, until now. On page 233 of this issue, Walter, Wolk and Neuhäuser¹ report that one member of this population has finally been seen.

To be sure, a previous paper by Stocke et al.² had reported the detection of a *candidate* old isolated neutron star. So what makes the authors of the current paper confident enough to call their result an actual discovery? The answer rests on two pillars: a highly unusual spectrum, and a bound on the distance to the source.

It was first suggested a quarter of a century ago^{3,4} that fading, old, isolated neutron stars can be reincarnated as faint X-ray sources by accreting gas directly from their surrounding medium. The radiation output from these objects is virtually all in soft X-rays, with a quasi-thermal spectrum at a temperature of about 10^6 K. Being faint they can only be seen (with current detectors) within a few hundred parsecs (pc) of the Sun. The source that Walter et al. found appears to have all the right properties: it is an unidentified soft X-ray source (detected by Rosat, the Röntgen X-ray satellite) with no apparent counterpart at any other wavelength. The authors put a lower limit of 7,000 on the X-ray to optical flux ratio, an incredibly high value which rules out virtually all other types of astronomical object from white dwarfs to active galaxies. In addition, this source is in line with a molecular cloud 120 pc away. Because there appears to be little photoelectric absorption of soft X-rays by intervening material, the authors conclude that the source must be in front of the cloud, putting it squarely in the solar neighbourhood and ruling out any extragalactic origin.

Assuming these pillars hold up, a concern immediately arises: why have there been so few detections (only two at most)? Both the Stocke and Walter sources were found serendipitously; systematic searches among, for example, unidentified objects in the Rosat database, have not yet yielded any 'hits' (R. Danner, personal communication). This is despite predictions in several theoretical papers^{5–7} that there should be tens to thousands of old, isolated neutron stars screaming for detection. Astronomers have scrambled for explanations.

There are many possible reasons why the predictions are off. It could be that (1) the neutron star birth rate has been grossly overestimated; (2) they never manage to accrete from their surroundings, either because their rotation does not slow down sufficiently or because their magnetic fields remain too high (in either case infalling material would be dragged around and slung away from the star); (3) they are indeed able to accrete from the interstellar medium but the accretion heats up the gas which in turn quenches the accretion; (4) most neutron stars are born with a high kick velocity, for which accretion is relatively inefficient; or (5) the spectrum from most neutron stars is non-thermal, and most of their emission is at higher energies than the Rosat spectral band. Given the current understanding of supernova rates, pulsar birth rates and Galactic chemical evolution, most astronomers would be loath to adopt item 1. That leaves the next four possibilities (at least), which are all variations on the same theme, namely that astronomers currently have a very poor understanding



Distribution on the sky of 497 radio pulsars (young neutron stars). The Galactic plane is at zero latitude ($b=0$), and the Galactic Centre is in the middle of the diagram. The distribution of old, isolated neutron stars is not known.

of the long-term evolution of neutron stars and their interaction with surrounding media. This ignorance can be forgiven because they have had no reincarnated objects from the cosmic graveyard to guide their studies.

What is sorely needed now is a confirmation of the nature of this source. Despite the evidence, scepticism remains. It is disputed as to how well the authors have measured soft X-ray absorption, and that could significantly affect the distance bound. How well they have fixed the position of their source is also in question, and this could affect their conclusion about the absence of counterparts at other wavelengths. Difficult as it will be, spectroscopy may be the final arbiter. If this object is indeed sitting in front of the molecular cloud and accreting from the low-density interstellar medium, it should be surrounded by an ionization nebula, with its own distinctive set of line signatures in the optical^{4,6,8} and the infrared. If the neutron star has remained strongly mag-

netized, with a field strength $\sim 10^{12}$ gauss, there should be a cyclotron emission feature in the hard X-ray band at ~ 10 keV, on top of the soft X-ray continuum⁹.

The final judgement is not in on this source, but its case remains the most compelling to date. This direct evidence for a vast stellar graveyard supports the long held belief that massive stars have been forming and dying ever since our Galaxy came into existence. Instead of theorizing in the dark, astronomers can finally place the study of the long-term evolution of the superdense remnants on a firm observational footing. With more detections, they can also hope to understand better how the interstellar medium of our Galaxy has been enriched over the aeons with all the elements heavier than hydrogen and helium. It is from this enriched medium that subsequent generations of stars and planets — including the Sun and the Earth — evolved.

Whether it be this source or some other, given the current observational and

theoretical efforts, it seems only a matter of time before some member of the graveyard's quiet majority is identified. Such an event would usher in a new era in the study of neutron stars and of Galactic chemical evolution. But what if the quiet majority is actually a silent majority? □

John C. L. Wang is in the Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA.

1. Walter, F. M., Wolk, S. J. & Neuhauser, R. *Nature* **379**, 233–235 (1996).
2. Stocke, J. T., Wang, Q. D., Perlman, E. S., Donahue, M. E. & Schachter, J. F. *Astr. J.* **109**, 1199–1203 (1996).
3. Ostriker, J. P., Rees, M. J. & Silk, J. *Astrophys. Lett.* **6**, 179–184 (1970).
4. Shvartsman, V. F. *Soviet Astron.-AJ* **14**, 662–668 (1971).
5. Treves, A. & Colpi, M. *Astr. Astrophys.* **241**, 107–111 (1991).
6. Blaes, O. & Madau, P. *Astrophys. J.* **403**, 690–705 (1993).
7. Madau, P. & Blaes, O. *Astrophys. J.* **423**, 748–752 (1994).
8. Blaes, O., Warren, O. & Madau, P. *Astrophys. J.* **454**, 370–381 (1995).
9. Nelson, R. W., Wang, J. C. L., Salpeter, E. E. & Wasserman, I. *Astrophys. J.* **438**, L99–L102 (1995).

ANIMAL BEHAVIOUR

The question of culture

Christophe Boesch

FOR centuries, the manufacture and use of tools have been taken as the hallmark of humanity. Manufacture, especially, has been considered to be the defining criterion, in that it requires the artificial imposition of a certain shape on an object that might have looked very different in its original form.

On the face of it, then, Gavin Hunt's report on page 249 of this issue¹ of precisely such performances in wild crows tremendously expands the boundaries of human characteristics. Hunt made his observations of the birds in New Caledonia, an island group some 900 miles off the northeast coast of Australia. He collected many discarded tools (both twigs with hooks on and stepped-cut leaves, used to tease out invertebrate prey from holes), and recorded tens of instances of actual tool use and a few examples of tool manufacture.

All current theories about the evolution of mankind rely on tool manufacture and use as being central behaviours that distinguished our early ancestors from apes. There is no evidence of tool use by species of *Australopithecus*^{2,3}, and the appearance of tools seems to have coincided with the emergence of *Homo habilis*, some 2 million years ago. Along with enlarged brain capacity, tool use is the most obviously distinctive trait of the genus *Homo*. Nonetheless, the tools associated with *H. habilis* are still quite crude and are not systematically standardized in shape and size. Only when

H. erectus appeared, about 1.8 million years ago, did standardization and sophistication in the form of the tools become the rule (for example in the shape of hand axes, which remained stable for about a million years). This imposition of regularity of form following a mental representation of the tool maker has been seen as the hallmark of culture, and it is an argument that has also been applied to other animals using tools, especially chimpanzees^{4,5}. The issues, however, are not straightforward.

Hunt suggests that the New Caledonian crows demonstrate three features new to tool-use in free-living, nonhuman animals, and that these features appeared first among early humans during the Upper Palaeolithic. The features are the use of hooks, the production of distinct tool types and a high degree of tool standardization. The crows were observed to use hooked twigs in hunting small invertebrates, and the intentional making of such hooks, as suggested by Hunt, would say a great deal about the crows' resourcefulness. Animal use of hooks has been observed only rarely, the most dramatic example that I know of coming from chimpanzees in Bossou (Guinea) which employed long, hooked branches in order to reach a branch of a fig tree that they could not reach on their own from below⁶.

We should expect distinct tool types to be produced, when a particular technique works, only if the tool fulfils some specific

criteria. For example, digging insects out of a 10-cm-deep hole can be performed only with an elongated object at least 10 cm long and a diameter smaller than that of the hole. Shorter or wider tools do not work, and by trial-and-error learning the tool user can produce distinct tools for this task. Chimpanzees are known to produce task-specific tools that vary distinctly in material, size, length, hardness and weight^{4,5}. It is intriguing to see that this might also be the case for the New Caledonian crows, but this aspect of a tool can be more a reflection of the requirement of the tasks performed than a sign of high technical capability.

Evidence of standardization in tool production would be a more decisive step in deciding whether such a capability is indeed being manifested. Standardization can result from two processes: first, an object may be progressively modified until the user can employ it for an intended purpose; or, second, the object may be completely modified for the purpose before use. In the first process a trial-and-error approach, requiring no idea about the shape of the tool, is sufficient. For example, if an animal wants to reach insects within a hole in a tree, a leafy twig does not go deep enough into the hole. It is only by progressively removing the leaves and other parts that the object will become a proficient tool. In the second process, the animal makes all the necessary changes beforehand and it is this process that is considered by some to be the characteristic of the existence of culture.

It is, however, extremely difficult to differentiate between the two processes on the basis of tools collected on the ground after use. Hunt collected 305 tools, but

saw only four instances of tool manufacture. In all four cases manufacture was completed before use, but it is risky to generalize from such a limited sample. By contrast, in one chimpanzee population that I have studied, in 93.5 per cent of a total of 363 sticks modified as tools all of the necessary modifications were performed before use⁵. In this case, standardization resulted from the second process and we can say with some certainty that chimpanzees have a preconceived idea of what shape an object must have to become a tool.

There is also more to the question of whether standardization of tools can be said to be a mark of the existence of culture. Hitherto in this article, the term standardization has been discussed as the result of task requirement combined with sophisticated mental abilities. But the existence of culture as such would also require that the standardization observed is underpinned by social as well as ecological criteria. So there is a third dimension to standardization: in free-living animals we would need evidence of a task that could be solved with at least two different possible types of tool but only one type is used in a given social group. In this case, part of the standardization would result from the tool users following social models. As far as I know, this has been demonstrated only in human beings and, again, chimpanzees — different populations of chimpanzees have been observed to 'dip' for the same species of ants by using sticks of different but discrete sizes, all members of one population producing tools of a single type⁵.

All in all, Hunt's fascinating paper gives much food for thought and argument. Although tool manufacture and use have been observed in many animal species⁴, at the least the work to date on the New Caledonian crows shows that tool use in birds is less stereotyped than previously thought. We can hope that further details will follow in due course. For instance, it would be especially interesting to know how much of the tool behaviour is anticipated and is truly a cultural product of crow society; if, to any extent, this is the case, it would add an important element to our knowledge of culture among animals. □

Christophe Boesch is at the Institute of Zoology, University of Basel, Rheinsprung 9, 4051 Basel, Switzerland.

New trick from an old foe

Bryan R. Cullen

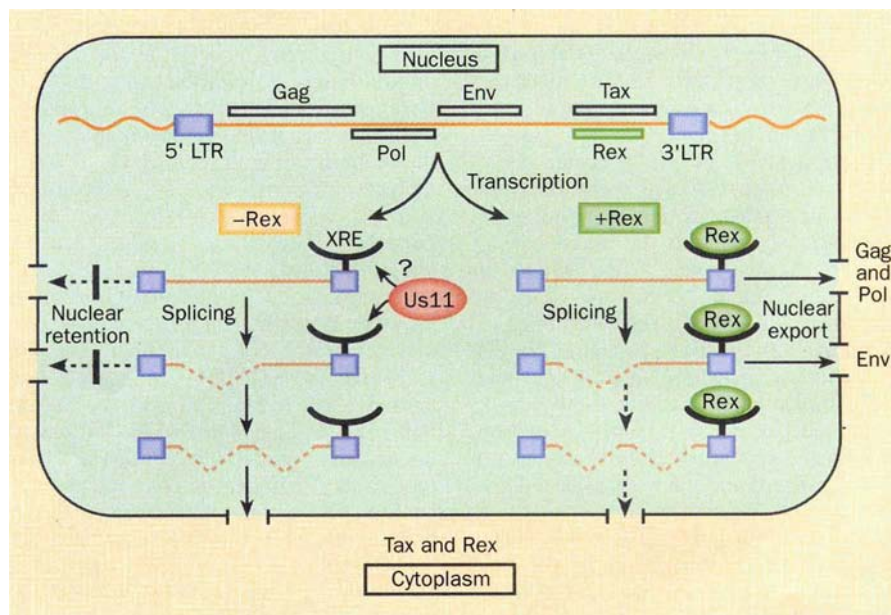
THE complex human retroviruses human T-cell leukaemia virus type-I (HTLV-I) and human immunodeficiency virus type 1 (HIV-1) regulate the order of their own gene expression so that regulatory proteins are made early after a cell has been infected while structural gene expression is delayed¹. This temporal regulation was unprecedented among retroviruses at the time of its discovery, but it has turned out to be characteristic of DNA tumour viruses, including all herpes viruses.

However, whereas the order of gene expression is transcriptionally regulated in DNA tumour viruses, it is regulated entirely at the post-transcriptional level in HTLV-I and HIV-1 (ref. 1). A study by Diaz *et al.* (page 273 of this issue²) shows that the Us11 gene product of herpes simplex virus type 1 (HSV-1) can also post-transcriptionally activate the expression of HIV-1 and HTLV-I envelope structural protein, and thus raises the possibility that this level of regulation is not unique to complex retroviruses.

Transcription of chromosomally integrated HTLV-I and HIV-1 proviruses starts from a promoter located in the 5' long terminal repeat (LTR) and ends in the 3' LTR to give a genome-length tran-

script that also serves as the messenger RNA for synthesis of the Gag and Pol structural proteins (see figure). In HTLV-I, *env* mRNA is derived by splicing out an intron consisting of *gag* and *pol* sequences while removal of a second intron overlapping *env* generates a bicistronic mRNA encoding the Tax and Rex regulatory proteins.

Because HTLV-I replication requires the expression of unspliced, singly spliced and fully spliced transcripts, processing of HTLV-I RNAs is designed to be inefficient and incomplete. However, early after infection these incompletely spliced HTLV-I mRNAs are sequestered in the infected cell nucleus and thus are not translated³. Although the reason for this retention remains uncertain, the presence of unused splice sites may induce nuclear retention by cellular factors, termed 'commitment factors', which normally prevent the nuclear export and translation of incompletely spliced cellular mRNAs⁴. Here, their action would block the expression of HTLV-I transcripts encoding the viral structural proteins, but would not interfere with the nuclear export of the fully spliced RNA encoding Tax and Rex.



HTLV-I is a retrovirus — an RNA virus that replicates via a DNA intermediate. Upon infection of a cell, the virus's reverse transcriptase makes a DNA copy of the viral RNA genome, which is integrated into the cell's chromosomal DNA as a provirus. The virus then uses the host's machinery to transcribe its own DNA as one long pre-messenger RNA encoding a number of different gene products. Unspliced or singly spliced 'late' RNAs encode the structural proteins Gag, Pol and Env; a fully spliced 'early' RNA encodes the regulatory proteins Tax and Rex, which must be made early in order to enter the nucleus and increase viral expression. In the absence of the essential Rex protein, the messenger RNAs that encode the viral structural proteins are retained in the nucleus and, hence, not expressed. Rex induces their nuclear export and expression after first binding to the viral Rex response element (XRE) RNA target. The HSV-1 Us11 protein has now been shown also to bind the XRE and to activate the cytoplasmic expression of the late *env* mRNA. LTR, long terminal repeat.

- Hunt, G. R. *Nature* **379**, 249–251 (1996).
- Jones, S. *et al.* *The Cambridge Encyclopedia of Human Evolution* (Cambridge Univ. Press, 1992).
- Leakey, M., Feibel, C. S., McDougall, I. & Walker, A. *Nature* **376**, 565–571 (1995).
- Beck, B. B. *Animal Tool Behaviour: The Use and Manufacture of Tools by Animals* (Garland, New York, 1980).
- Boesch, C. & Boesch, H. *Folia Primatol.* **54**, 86–99 (1990).
- Sugiyama, Y. & Koman, J. *Primates* **20**, 513–524 (1979).

Subsequently, the accumulating Rex protein induces the sequence-specific nuclear export of these sequestered viral RNAs after first binding to, and multimerizing on, the *cis*-acting Rex response element (XRE) RNA stem-loop structure¹. Similarly, the HIV-1 Rev protein induces the nuclear export of incompletely spliced HIV-1 transcripts after binding to the functionally equivalent Rev response element (RRE). In addition to domains required for RNA binding and multimerization, Rex and Rev contain a critical, leucine-rich, activation domain that is recognized as a potent nuclear export signal by the host cell⁵. The nucleocytoplasmic transport of target RNA species induced by Rex and Rev therefore results from the assembly of multiple nuclear export signals onto these RNAs.

Diaz and co-workers² now show that the Us11 protein encoded by HSV-1 can functionally substitute for HTLV-I Rex, and to a lesser extent for HIV-1 Rev, in inducing the expression of viral envelope proteins. In the case of HTLV-I, this expression reflects the cytoplasmic appearance of incompletely spliced *env* mRNA, whereas an incompletely spliced cellular mRNA remains entirely in the nucleus. The Us11 protein binds to both the XRE and the RRE *in vitro*, and induction of HTLV-I envelope protein expression by Us11 requires the presence of the XRE in *cis*. Overall, these data show that the HSV-1 Us11 protein can substitute for Rex in inducing the expression of a late HTLV-I protein and suggest that Us11 is doing so by a similar mechanism.

Does this result imply that HSV-1 is likely to facilitate HTLV-I or HIV-1 late gene expression in coinfecting cells? Two considerations suggest that this is unlikely. First, HSV-1 grows preferentially in epithelial and neuronal cells whereas HTLV-I and HIV-1 prefer lymphoid or (for HIV-1) myeloid cells. Dual infections are therefore likely to be rare. Second, HSV-1 inhibits host-cell gene expression soon after infection. Because human retroviruses rely extensively on the host cell to express their genetic material, HSV-1 probably significantly inhibits HTLV-I and HIV-1 gene expression when dual infection does occur.

If Us11 is functionally comparable to Rex and Rev, the next question is whether it has a similar role in HSV-1 replication. Us11 is expressed late in the HSV-1 life cycle, is packaged into HSV-1 virions at about 800 molecules per virion and is then released into the cytoplasm of cells shortly after infection⁶. By analogy to the VP16 regulatory protein, which is also delivered to cells as part of the HSV-1 virion, it has been proposed that Us11 acts early after infection.

However, mutational inactivation of

the Us11 gene only minimally affects HSV-1 replication in culture. Indeed, the only known function of Us11 is to inhibit the appearance of a truncated, non-polyadenylated HSV-1 transcript of unknown significance termed $\Delta 34$ RNA (ref. 7). Although it has been suggested that Us11 may reduce the synthesis of $\Delta 34$ RNA by blocking premature termination of transcription of the UL34 gene⁷, it is conceivable that it instead inhibits the endonucleolytic cleavage of full-length UL34 RNA within the cell nucleus by inducing its rapid export. Although the effect of Us11 on the subcellular distribution of either the truncated $\Delta 34$ or the full-length UL34 RNA remains unknown, Us11 does bind to both RNA species specifically *in vitro*⁷.

The role of Us11 in the HSV-1 life cycle therefore remains unclear, but it shares the ability of Rev and Rex to bind to specific RNA sequences, to form multimers and to localize to the nucleolus (although Us11 differs from Rex and Rev in also associating with cytoplasmic ribosomes⁶). All three proteins are also highly basic, though Us11 lacks a typical arginine-rich RNA-binding motif and instead contains 24 carboxy-terminal repeats of the sequence Arg-X-Pro. An obvious, and potentially significant, difference is that Us11 does not contain any sequence

evidently similar to the conserved, leucine-rich nuclear export signal that is crucial in the mechanism of action of Rex and Rev.

So whereas Us11 can at least partly substitute for Rev and Rex in mediating the nuclear export and translation of mRNAs encoding retroviral structural proteins, it remains unclear whether it is truly acting through the same mechanism. But if Us11 does contain a novel nuclear export signal that can promote the transfer of specific retroviral and (presumably) HSV-1 transcripts, this would represent an exciting and surprising dissemination of what was previously perceived as a unique retroviral regulatory pathway. □

Bryan R. Cullen is in the Howard Hughes Medical Institute and Department of Genetics, Duke University Medical Center, Room 426 CARL Building, Research Drive, Durham, North Carolina 27710, USA.

1. Cullen, B. R. *Microbiol. Rev.* **56**, 375–394 (1992).
2. Diaz, J. J. *et al. Nature* **379**, 273–277 (1996).
3. Hanly, S. M. *et al. Genes Dev.* **3**, 1534–1544 (1989).
4. Legrain, P. & Rosbash, M. *Cell* **57**, 573–583 (1989).
5. Fischer, U., Huber, J., Boelens, W. C., Mattaj, J. W. & Lührmann, R. *Cell* **82**, 475–483 (1995).
6. Roller, R. J. & Roizman, B. *J. Virol.* **66**, 3624–3632 (1992).
7. Roller, R. J. & Roizman, B. *J. Virol.* **65**, 5873–5879 (1991).

SUNSPOTS

Something stirs under the Sun

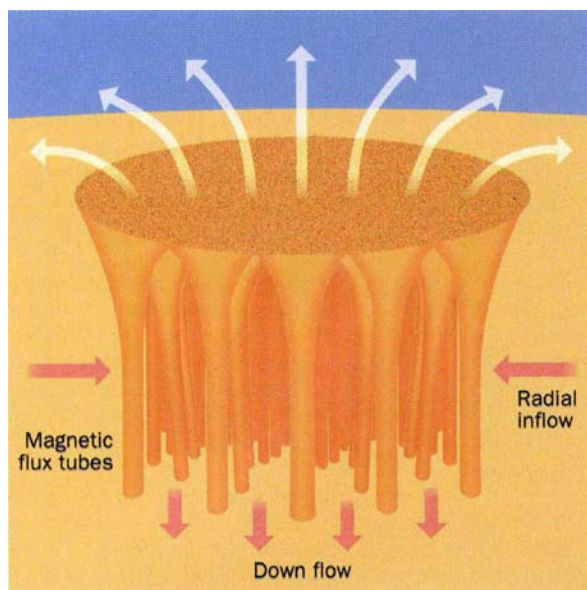
Eugene Parker

THE naked eye can occasionally make out dark spots on the rising or setting Sun in a murky sky. Oriental astronomers noticed this phenomenon two thousand years ago and built it into their astrological system, and spots were 'discovered' in the West in 1611 with the invention of the telescope. George Ellery Hale first measured their intense magnetic fields¹ (2–3 kilogauss, about as strong as the field in the gap of the most powerful permanent magnet) in 1908, and the physics of the sunspot has been pursued by both observers and theoreticians since then. Yet it is not possible today to explain why a star like the Sun is compelled by the laws of nature to produce cool concentrated magnetic patches on its surface. However, a new and ingenious observational technique called helioseismic tomography has been developed to probe the fluid motions below the visible surface, and with its first trial application (reported on page 235 of this issue²) Duvall *et al.* have established that there is a strong down draught beneath a sunspot.

On close scrutiny, the light from other stars shows regular time variations that can be explained only by dark patches

rotating around with the star. In some extreme cases (young, low-mass stars) these patches may cover as much as half of the visible disk, so there is nothing special about dark spots on a star. The ubiquitous 'star spot' is a natural consequence of convection and magnetic fields below the surface. The problem is that we do not understand the combined effects well enough to explain it theoretically. Observations are needed, but the problem is that one can see only the surface of the spot — a cool area (3,900 K) against the hotter (5,600 K) background of the normal photosphere.

It was pointed out decades ago³ that the magnetic field is so strong as to stifle the convection responsible for carrying heat to the surface, hence the coolness. However, the magnetic field exerts an enormous pressure, about 4×10^5 dyn cm⁻², which must be confined by the surrounding gas if the field is not to expand and disperse. The answer can only be that the magnetic field is rammed into the spot by a powerful converging flow⁴, which then forms a down draught underneath. No such flow is visible, so it must lie somewhere below the surface convective cells.



The subsurface structure of a sunspot, with the magnetic field extending upwards and outwards from the visible surface. The short, straight arrows indicate the fluid motion deduced from tomography, flowing inwards and then downwards between the separate magnetic flux bundles.

Duvall *et al.* report the first observational look at the fluid motions beneath the visible surface. Their analysis concentrates on an active region of the Sun, and shows a downward flow travelling at 2 km s^{-1} to a depth of 2,000 km beneath a sunspot, implying the radial inflow predicted by theory. The extension of the inflow to the central region under the spot implies that the magnetic field at depths of $\sim 1,000 \text{ km}$ is separated into many intense flux bundles⁵, to allow the inflow to stream through the relatively field-free interstices (see figure).

The downflow was detected by looking at solar vibrations. The continual turbulent convection within the Sun is a source of sound waves, exciting thousands of modes of oscillation. Conventional helioseismology studies the density, temperature, mean molecular weight, and rate of rotation of the Sun as a function of depth by precise measurement of the frequencies of these modes, but it cannot give information about the local fluid motion. In contrast, helioseismic tomography can discern the fluid motion by measuring the correlation of acoustic motions at many pairs of points on the surface of the Sun. The ray path and transit time by which a sound wave propagates from one point to another can be calculated from the known temperature and molecular weight. A delay indicates a head wind somewhere along the ray path, and an early arrival indicates a tail wind.

The effect of a magnetic field is somewhat more complicated, as it increases the sound speed perpendicular to the field but not parallel. Using a computer to compare the thousands of pairings between many points allows construction of a map

of fluid velocity as a function of depth below the surface, as well as providing basic information on the form and strength of the mean magnetic field.

Sunspot dynamics are beginning to come under the direct eye of observation. The authors emphasize that this initial breakthrough is only the first exploratory application of tomographic analysis to the activity within the Sun. Observations are possible at substantially higher resolution and to greater depth. Such observations are needed, for the convection beneath the surface of the Sun is a magnetohydrodynamic jungle. It is conjectured that the complicated magnetic field observed at the surface of the Sun is an indication that the flux tubes (or 'fibrils') are in a

complex tangled state throughout much of the convective zone, extending to a depth of about $2 \times 10^5 \text{ km}$. The field intensity, characteristic diameter and degree of entangling of the fibrils is not known, so there is no way to calculate the details of their dynamic interaction with each other and with the fluid. Yet the generation of the magnetic fields responsible for the sunspot cycle and for the associated 11-year variation in brightness depends critically on the mean motion and interaction of individual magnetic fibrils. It is their buoyancy and convective transport that are responsible for flares, extreme ultraviolet emission, X-rays, the corona and the solar wind.

Then when we reflect that the Sun is just an average main sequence star, it becomes clear that until we understand it, we understand little of other stars. There is also the practical need to understand the effects of solar variation on our climate. Helioseismic tomography is a promising step in probing the magnetic jungle that has so much to do with the various forms of energy output of the Sun. □

Eugene Parker is at the Enrico Fermi Institute, the Department of Physics and the Department of Astronomy, University of Chicago, Laboratory for Astrophysics and Space Research, 933 East 56th Street, Chicago, Illinois 60637, USA.

1. Hale, G. E. *Astrophys. J.* **28**, 100, 315 (1908).
2. Duvall, T. L. Jr, D'Silva, S., Jefferies, S. M., Harvey, J. W. & Schou, J. *Nature* **379**, 235–237 (1996).
3. Biermann, L. *Vierteljahrsschr. astr. Ges.* **76**, 194 (1941).
4. Meyer, F., Schmidt, H. U., Weiss, N. O. & Wilson, P. R. *Mon. Not. R. astr. Soc.* **169**, 35 (1974).
5. Parker, E. N. *Astrophys. J.* **230**, 905–913 (1979).

War studies

J. B. S. HALDANE once claimed that a democratic election predicts the result of a civil war without having to fight one. He was oversimplifying. A war, civil or otherwise, is a competition for territory between rival tribes. A true substitute for war would have to predict, not just the 'winner', but also the final boundary between the tribes. Even the recent Bosnian settlement had to reflect roughly the military gains of the competing tribes. Daedalus wants to determine such gains not by war, but by war gaming.

All professional armies these days conduct simulated computer exercises: not silly shoot 'em up arcade games, but complex, serious simulations. The best of them can model a battle on difficult terrain between forces with men and weapons of disparate strengths and weaknesses, and predict the likely outcome. Daedalus wants a really big modelling program of this sort to be developed, and put at the disposal of the United Nations.

International mediation would be transformed. Instead of its blundering insertion of token forces, the UN would use satellite intelligence and ground observers to assess the whole situation. It would run the simulation, and predict the outcome. It could then say to the two sides: "This is what will happen. Instead of expending vast amounts of blood and treasure proving it, why not accept the result, and save yourselves the trouble? Redraw the tribal boundaries from our computer output, exchange whatever populations would have been uprooted, and resume peaceful coexistence".

The biology of the proposal is sound; our cousins the apes do something like it already. Their rival tribes howl and threaten each other along their mutual territorial boundary, and the display predicts for both parties the result of a fight without the nuisance of having one. Computer war-simulation will do the same for us. Much development, however, will be needed to make it credible. For a start, it will have to replicate the early skirmishes before the UN was called in. Fortunately it need not be perfect. Much of the optimized final boundary will be insensitive to small errors: it will lie along stable, easily defended rivers, mountain ridges and so on. In any case, a small needless loss of territory is preferable to a crippling fight for a bit more. The big power blocs of the Cold War world are now breaking up into ever more squabbling tribes, each demanding its own territory and certain to be a permanent nuisance till it gets it. Daedalus's computer war-simulator should minimize the misery of the transition.

David Jones

Pterodaustro's true teeth

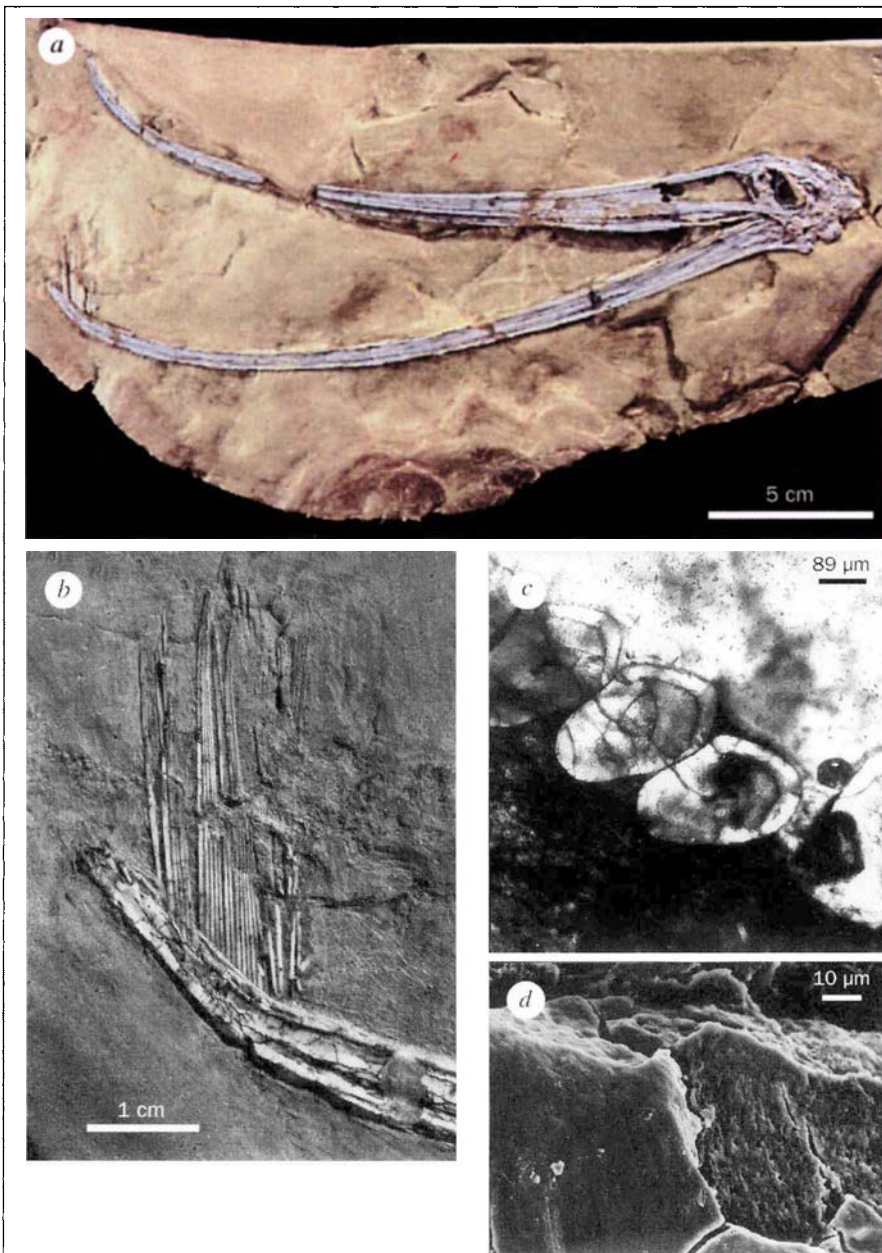
SIR — With a wing span of 1–2.5 metres, a long, thin, upward-curved skull, and jaws bearing up to 1,000 filament-like teeth, the filter-feeding, Early Cretaceous *Pterodaustro guinazui*^{1,2} is among the most peculiar pterosaurs ever found (*a, b* in the figure). *Pterodaustro's* bizarre mandibular teeth have long puzzled palaeontologists. The fine structure and baleen-like arrangement of the teeth (*a, b* in the figure) led some to regard them as keratin-like or bristle-like structures that were not actual teeth^{3,4}. Our study, based on new specimens from laminated shales of the Lagarcito Formation of central Argentina⁵, provides a description of the intriguing filter-feeding apparatus of *Pterodaustro* and confirms its dental nature.

The mandibular teeth of *Pterodaustro* are oval to sub-elliptic (*c, d* in the figure). Towards the tip of the jaw, the teeth are much thicker than the middle and caudal ones. Throughout the jaw, the teeth are singly arranged and close together (see figure). For most of the jaw's length they are set in a groove. Near the tip of the dentary, however, the teeth are set in individual, shallow alveoli. The dentary wall lateral to the tooth row is one-third to one-quarter thinner than the medial wall. The medial wall is also much higher than the lateral wall. The external surfaces of the teeth are smooth. The premaxillary and maxillary dentition of *Pterodaustro* is formed by hundreds of tiny teeth of similar size, with spatulate crowns and slender, conical bases. These teeth are not set in alveoli, but were probably embedded in a corneous structure or a muscular 'lip'.

Teeth from the middle section of a mandible of *Pterodaustro* were sectioned both longitudinally and transversely (*c, d* in the figure). In one transverse section, at least ten teeth were visible, with seven of them in an excellent situation for detailed measurement. The cross-sections revealed a thin, peripheral layer of enamel, indicated to be non-prismatic by both light and scanning electron microscopy

(*c, d* in the figure). The rest of the tooth consists of dentine with a more or less central pulp cavity. Incremental lines were not observed in either the enamel or the dentine. The dentinal tubules are arranged radially around the pulp cavity. Transverse

and longitudinal sections of the teeth indicate that the dentinal thickness is variable around the pulp cavity. Scanning electron microscopy provided a topographical view of the dentine and revealed the dentinal canals (*c, d* in the figure). In some regions, casts of odontoblast processes with multiple lateral branching were observed within the dentinal canals.



a, b, *Pterodaustro guinazui* (UNSL-GEO-V-57, Universidad Nacional de San Luis (Geología), Argentina) from the Early Cretaceous (Albian) Lagarcito Formation of central Argentina. *a*, Left lateral view of the skull. Most of the dentary teeth are missing; only those in the tip of the jaw are preserved. *b*, Detail of the mandibular tip (photographed under UV light). Note that both the left and right dental rows are preserved. The laminated shales of the Lagarcito Formation were deposited in distal sections of a large, perennial, shallow body of fresh water, dominated by warm and semi-arid climatic conditions⁵. *c, d*, Mid-dentary teeth of *Pterodaustro* (UNSL-GEO-V-55). *c*, Transverse section. Note the peripheral cap of enamel and the fine dentinal tubules arranged radially around the pulp cavity. Cross-sectional average tooth length is 318 μm (including the enamel thickness of $\sim 38 \mu\text{m}$, $n=7$), while the average width is 197 μm ($n=7$). *d*, Scanning electron micrograph of a naturally fractured tooth showing the aprismatic nature of the enamel and dentinal canals in the dentine. Note the scratch marks on the enamel in the lower left region. The sample was treated with 5% HCl for 30 s.

1. Bonaparte, J. *Acta geol. Iilloana* **10** (10), 207–226 (1971).
2. Sanchez, T. *Publins Mus. mun. Cien. nat. Mar del Plata* **1** (9), 263–272 (1973).
3. Wellnhofer, P. *The Illustrated Encyclopedia of Pterosaurs* (Crescent, New York, 1991).
4. Benton, M. *The Reign of the Reptiles* (Crescent, New York, 1990).
5. Chiappe, L. M. et al. in *Extended Abstracts, 2nd. int. Symp. Lithogr. Limestones 57–60* (Ediciones de la Universidad Autónoma de Madrid, Madrid, 1995).
6. Ellis, R. *The Book of Whales* (Knopf, New York, 1980).
7. Olson, S. L. & Feduccia, A. *Smithson. Contr. Zool.* **316**, 1–73 (1980).
8. Smithers, R. H. N. *The Mammals of the Southern African Subregion* (Univ. Pretoria Press, Pretoria, 1983).
9. Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).

The morphological and histological evidence presented here demonstrates that the remarkable mandibular filter-feeding apparatus of *Pterodaustro* is composed of true dental structures. Elaborate filter-feeding mechanisms are rare among adult tetrapods. The filter-feeding apparatus of *Pterodaustro* is not comparable to that of baleen whales⁶ or filter-feeding birds⁷, which are equipped with non-dental, epidermally derived sieving structures. Modified teeth for filtering are known for mammals (such as the crab-eater seal, *Lobodon carcinophagus*⁸), the Permian mesosaurs⁹, and other pterosaurs (such as *Ctenochasma* and *Gnathosaurus*³), although none of these shows the extraordinary specialization of *Pterodaustro*.

L. M. Chiappe

Department of Ornithology,
American Museum of Natural History,
New York, New York 10024, USA

A. Chinsamy

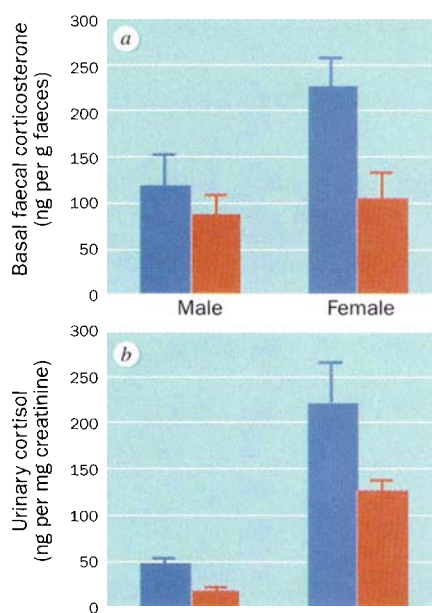
Earth Sciences Division,
South African Museum, PO Box 61,
Cape Town 8000, South Africa

Social stress and dominance

SIR—It has recently been shown that social dominance can carry subtle costs^{1–5}. Although it is generally thought that social stress falls more heavily on subordinates than on dominants, few studies have examined how stress relates to rank in the wild. We have investigated the levels of glucocorticoid stress hormones in two carnivores that live in complex societies, African wild dogs (*Lycaon pictus*) and dwarf mongooses (*Helogale parvula*). For these species, chronic stress is apparently a cost of social dominance rather than of subordination.

The link between social subordination and elevated glucocorticoids has primarily been shown in captivity. Although wild subordinate baboons have chronically elevated glucocorticoid levels⁶, data from other free-living species do not confirm that social subordinates experience greater stress than dominants^{7–9}. In the wild, subordinates can often avoid aggressive interactions, while dominants initiate aggression to reinforce their social status. Consequently, dominant dwarf mongooses and wild dogs engage in aggressive interactions more often than do subordinates, particularly during mating periods (wild dog females: $t_{33}=2.30$ (Student's *t*-test), $P=0.015$; wild dog males: $t_{78}=2.81$, $P=0.003$; dwarf mongooses, see ref. 10). High rates of aggression may be stressful, regardless of whether the encounters are won. If so, dominants would suffer greater social stress than subordinates.

We measured the urinary cortisol levels



c.v. was 3.4%. The inter-assay c.v. was 18.4%. HPLC showed a single peak of immunoreactivity in mongoose urine, which coeluted with free cortisol. Error bars show 1 s.e.

of dwarf mongooses (Serengeti National Park, Tanzania; 97 females, 82 males, 740 samples) and the faecal corticosterone levels of wild dogs (Selous Game Reserve, Tanzania; 22 females, 34 males, 216 samples). In both studies, samples were collected noninvasively from known individuals whose social rank was determined from aggressive interactions¹⁰. For African wild dogs, dominants had higher basal corticosterone levels than subordinates (*a* in the figure; partial $F_{1,207}=6.76$, $P=0.01$, controlling for gender and reproductive condition). Because we collected samples from unhandled animals, we have no data on corticosteroid responses to acute stress in wild dogs. For female dwarf mongooses, basal cortisol levels were higher in dominants than in subordinates (*b* in the figure; $t_{99}=6.37$, $P<0.001$). Nonetheless, dominant females produced higher cortisol levels than subordinates in response to the stress of trapping (*b* in the figure: $t_{74}=3.06$, $P=0.003$), indicating that their acute stress response was not compromised by chronically elevated cortisol. In male mongooses, basal cortisol levels did not correlate with rank, but dominant males showed low peak cortisol levels in response to trap-stress ($t_{91}=3.01$, $P=0.003$), indicating that their acute stress response was compromised.

Because chronically elevated glucocorticoids can inhibit reproduction^{6,7}, it is intriguing that dominants are typically the only breeders within packs of both species (for example, dwarf mongooses, ref. 11; wild dogs, S.C. and N.M.C., unpublished data). Dominant wild dogs and dwarf mongooses apparently experience greater social stress than do subordinates. This result is predicted if aggression itself is stressful, because dominant dwarf

Glucocorticoid levels in African wild dogs and dwarf mongooses. *a*, Dominant African wild dogs (blue bars) show higher basal corticosterone levels than subordinates (red bars) ($F=6.76$; $P=0.01$). Corticosterone values are expressed as ng per g dry faeces, measured with a double-antibody ¹²⁵I-corticosterone radioimmunoassay (ICN Biomedicals). Assay sensitivity was 25 ng corticosterone per ml. The inter-assay coefficient of variation (c.v.) was 8.0% for a high-corticosterone control and 13.7% for a low-corticosterone control; the intra-assay c.v. was <10% for both controls. HPLC showed a single peak of corticosterone immunoreactivity in wild dog faeces, with polarity intermediate between cortisol and corticosterone. *b*, Dominant female dwarf mongooses show higher basal cortisol levels than subordinates ($t=6.37$; $P<0.001$) and higher peak cortisol levels in response to stress ($t=3.06$; $P=0.003$). The effect of rank on cortisol levels remains significant after controlling for the effect of diurnal variation. Reproductive condition (mating/non-mating) did not affect cortisol levels. Cortisol levels are expressed as ng of hormone per mg of urinary creatinine to control for urine concentration. Urinary cortisol was assayed at 1:250 dilution (0.5 ml) using a ³H-cortisol RIA kit (RSL Immuchem). The intra-assay

mongooses and wild dogs often assert their dominance through aggression. In contrast to our findings, dominant baboons rarely assert their rank⁴, and social stress falls on subordinates.

Social stress may vary with social conditions, so that dominants are stressed (rather than subordinates) if they are highly aggressive. In captivity, stress may fall on subordinates because they cannot avoid aggressive dominants. Broader data are needed to resolve the differences among species (including humans) and social contexts. Nonetheless, our findings show that social stress can be a cost that offsets the benefits of dominance.

Scott Creel

Nancy Marusha Creel

Field Research Center for Ecology
and Ethology,

Rockefeller University,
Millbrook, New York 12545, USA

Steven L. Monfort

Department of Reproductive Physiology,
Conservation and Research Center,
National Zoological Park,
Forest Royal, Virginia 22630, USA

1. Packer, C. et al. *Nature* **373**, 60–63 (1995).
2. Wasser, S. *Nature* **376**, 219–220 (1995).
3. Packer, C. *Nature* **377**, 689–690 (1995).
4. Altmann, J. et al. *Nature* **377**, 688–689 (1995).
5. Frank, L. G. et al. *Nature* **377**, 584–585 (1995).
6. Sapolsky, R. M. *Endocrinology* **116**, 2273–2278 (1985).
7. Wingfield, J. C. et al. *J. Zool., Lond.* **225**, 43–58 (1991).
8. Schoech, R. L. et al. *Condor* **93**, 354–364 (1991).
9. Mays, N. A. et al. *Auk* **108**, 619–637 (1991).
10. Creel, S. R. et al. *Anim. Behav.* **43**, 231–245 (1992).
11. Keane, B. et al. *Anim. Behav.* **47**, 65–75 (1994).

Scientific Correspondence

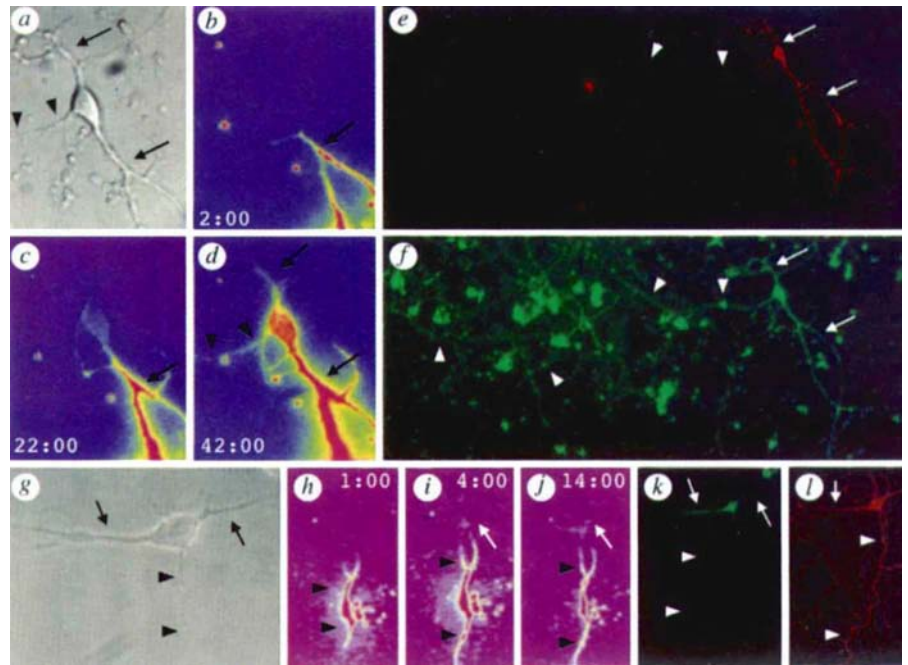
Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words.

No diffusion barrier at axon hillock

SIR — In epithelial cells, the tight junction poses a barrier preventing diffusional mixing of membrane proteins and lipids between apical and basolateral domains^{1,2}. Evidence for a functional equivalent to this tight junction at the axon hillock of neurons has been presented³. The viral protein haemagglutinin, which is restricted to axons, was used to induce the fusion of liposomes containing fluorescent phospholipids exclusively along axons. No decrease in staining in the proximal axon and no soma staining was observed, suggesting a diffusion barrier for lipids at the axonal hillock. It has been suggested, however, that the small proportion of label (~10%) that would diffuse onto the cell body during the experiment would not have been detected⁴. Here we present findings that argue against a lipid diffusion barrier at the neuronal axon hillock.

Applying a local perfusion labelling technique⁵, we used the fluorescent lipid DiI(C₁₂) to label a short segment of dendritic or axonal membrane of differentiated hippocampal neurons, near the soma, and followed the spread of fluorescence over time. DiI was seen to diffuse across the axon hillock from both somatodendritic (13/13 cases) and axonal sides (22/24 cases) within minutes (*b–d* and *h–j* in the figure). This observation was possible because of the very bright labelling given by DiI and the sensitive fluorescence imaging method used. In addition, labelling of dendrites as well as the axon was possible and so diffusion of label from both sides across the hillock could be followed. Labelling of dendrites resulted in a bright signal whose spread could be detected easily (*d* in the figure) but, as expected, soma labelling was faint after DiI incorporation into the axon (*j*). The fluorescence intensity showed smooth profiles after diffusion, with no abrupt changes at the axon hillock, consistent with free diffusion across this region.

Morphological criteria⁶ were used to identify axonal and dendritic processes before DiI labelling. This identity was subsequently confirmed by immunocytochemical staining against either a subtype of the glutamate receptor (GluR1) or microtubule-associated protein 2 (MAP2), both of which are restricted to somatodendritic domains^{6,7}. Preferential localization of the membrane protein GluR1 was observed in the somatodendritic domain, whereas a fluorescent lectin, succinylated concanavalin A, uniformly stained the entire neuronal surface (*f* in the figure). Thus, differential membrane protein distribution was maintained in those cells that had shown no diffusion barrier for DiI at the axon



Diffusion of the fluorescent lipid DiI across the axonal hillock. *a–d*, A 2-week-old neuron was labelled with DiI on a dendrite (axon marked by arrowheads, dendrites by arrows) and the spread of fluorescence followed over time (indicated by numbers in minutes) with a cooled CCD camera. Note the labelling of the soma by 22 min and of the proximal axon by 42 min. Immunocytochemical staining of the neuron showed preferential restriction of GluR1 to the somatodendritic membrane (*e*) and uniform distribution of surface receptors (*f*) for FITC-succinylated concanavalin A along all processes. *g–j*, A 2-week-old hippocampal neuron in culture was labelled with DiI on the axon and the spread of label followed over time. The fluorescence was detected at the soma (white arrow) 4 min after labelling. The axonal identity of the labelled process was confirmed by immunocytochemical staining against MAP2 (*k*) and tubulin (*l*; polyclonal antibody kindly provided by F. Solomon). We note that GluR1 staining showed a tapered gradient into the proximal axon, similar to that observed for MAP2 (ref. 7). Preparation of cultures and fluorescence imaging followed those reported previously^{5–7}. The fluorescence intensity was encoded by a continuous scale of pseudocolours. In *b–d* and *h–j*, low intensity is represented by purple/blue, intermediate intensity by yellow and high intensity by red.

hillock. Similar preferential somatodendritic localization of the cytoskeletal protein MAP2 was also observed (*k* in the figure). All nerve processes were revealed by immunostaining against tubulin (*l*).

Several lines of evidence indicate that the DiI label was properly incorporated into the plasma membrane and that the spread across the axon hillock was due to lateral diffusion within the plasmalemma. The rate of fluorescence spread is consistent with lateral diffusion of membrane lipids (see also ref. 5). The absence of significant lipid internalization is indicated by the ring staining of the soma and the absence of punctate fluorescence within the neuron. The use of a polarized excitation light source⁸ confirmed the proper orientation of the majority of DiI fluorophores within the plasma membrane matrix. Finally, in several cases we observed the spread of DiI along a labelled process that was not connected to the nearest soma but turned or grew alongside the soma. No staining of the soma was found in these cases. The faint soma staining observed after local labelling thus did not result from DiI molecules released from the labelled segment or from extracellular DiI crystals that diffused to the soma through the aqueous phase rather than along the labelled process.

Preferential localization of GluR1 in the somatodendritic membrane in the absence of an apparent diffusion barrier at the axonal hillock suggests that other mechanisms are involved in maintaining axonal and somatodendritic domains⁹. The lateral mobility of membrane proteins in these neurons (B.W. and M-m.P., unpublished results) and the tapered gradient of GluR1 staining into the proximal axonal domain (see figure) support the notion that steady-state asymmetry in the distribution of plasma membrane proteins may result from their selective insertion into and/or removal from the plasmalemma.

Bettina Winckler
Mu-ming Poo

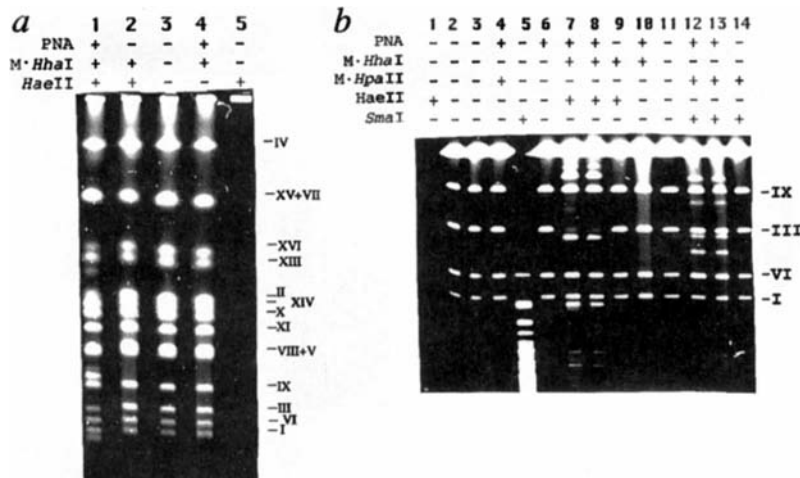
Department of Biological Sciences,
Columbia University,
New York, New York 10027, USA

1. Dragsten, P. R., Blumenthal, R. & Handler, J. S. *Nature* **294**, 718–722 (1981).
2. Gumbiner, B. *Curr. Opin. Cell Biol.* **2**, 881–887 (1990).
3. Kobayashi, T., Storrie, B., Simons, K. & Dotti, C. G. *Nature* **359**, 647–650 (1992).
4. Futerman, A. H., Khanin, R. & Segel, L. A. *Nature* **362**, 119 (1993).
5. Popov, S., Brown, A. & Poo, M-m. *Science* **259**, 244–246 (1993).
6. Caceres, A., Banker, G. A. & Binder, L. J. *J. Neurosci.* **6**, 714–722 (1986).
7. Craig, A. M., Blackstone, C. D., Huganir, R. L. & Banker, G. *Neuron* **10**, 1055–1068 (1993).
8. Axelrod, D. *Biophys. J.* **26**, 557–574 (1979).
9. McCloskey, M. & Poo, M-m. *Int. Rev. Cytol.* **87**, 19–72 (1984).

To assess the use of bis-PNA in cutting genomic DNA, we investigated cleavage of the 16 chromosomes of the yeast genome using two combinations of methylation and restriction enzymes: M-HhaI/HaeII and M-HpaII/SmaI (see

We have shown that short bis-PNA clamps used in conjunction with appropriate pairs of methylation and frequently cutting restriction enzymes can yield large

*To whom correspondence should be addressed.



Application of the proposed technique to the yeast genome. *a*, Results of pulsed-field gel electrophoresis (PFGE) separation of yeast genome (strain PSY316) after applying bis-PNA and the *M-HhaI*/*HaeIII* pair of enzymes. Yeast DNA in a low-melting agarose plug was equilibrated with buffer (20 mM MES, 2 mM EDTA, 10 mM NaCl, pH 6.3, 37 °C) and incubated with 0.2 μ M bis-PNA (10 h, 37 °C). Samples were equilibrated in buffer containing 25 mM sodium citrate and 5 mM β -mercaptoethanol (pH 7.0, 37 °C), and incubated overnight with 15 U *HhaI* methylase (4 °C, 100 mg ml⁻¹ BSA). S-adenosylmethionine (300 mM) was added and methylation performed (37 °C, 2.5 h). The reaction was stopped and the PNA-DNA complex was dissociated (50 min at 58 °C in 50 mM Tris-HCl, 500 mM NaCl, 10 mM EDTA pH 8.9 and 1% SDS). Samples were equilibrated in NEBuffer 4, digested with 4 U *HaeIII* restriction endonuclease (100 mg ml⁻¹ BSA, 4 h, 37 °C), and the reaction was then stopped (20 min, 50 mM EDTA, pH 8.7). Samples were equilibrated in TE buffer, pH 7.5, and loaded onto a 1% agarose gel. The PFGE was run on a CHEF Mapper system (Bio-Rad). *b*, Analysis using a PFGE regime permitting resolution of comparatively short DNAs (up to about 0.5 Mbp). Bis-PNA (0.4 μ M) and two pairs of methylation/restriction enzymes were used. Lanes 7, 8, results for the *M-HhaI*/*HaeIII* pair (20 and 25 U, respectively); lanes 12, 13, results for the *M-HpaII*/*SmaI* pair (10 and 15 U, respectively). Other lanes correspond to control experiments.

The invisible man of science

Michael Sherborne

HG: The History of Mr Wells. By Michael Foot. Doubleday: 1995. Pp. 318. £20.

FIFTY years after his death, H. G. Wells remains an anomaly among literary figures, first in having received a scientific education, second in having given science such a central place in his writings. Wells attended the Normal School of Science in London (later called the Royal College and later still Imperial College), where T. H. Huxley's lectures on biology and zoology made a deep impression on him. He then worked as a science teacher and journalist and wrote two biology textbooks, one of them with R. A. Gregory, an old college friend. When Gregory was appointed editor of *Nature*, Wells immediately became a contributor. He must have been delighted to appear in a journal for which Huxley had once written, yet already he had turned his own sights on quite a different career.

The success of his "scientific romance" *The Time Machine* (1895) established him as a creative writer of remarkable originality, with a keen interest in social issues and in what we would now call 'futurology'. In his new role, science remained extremely important to him. He relished the unfamiliar perspectives it opened on the human condition, its corrosive effect on superstition and prejudice, and the prospect that its discoveries would unleash radical social change. The best of his early science fiction stories — *The Time Machine*, *The War of the Worlds* and *The Island of Doctor Moreau* — not only draw on scientific ideas but also dramatize problems of ethics and evolution that had troubled Huxley. Are human beings, at bottom, any different from animals or machines? If, as Huxley claimed, our moral sense is based on the natural sympathy we feel for others of our kind, are we therefore justified in behaving unethically towards animals or so-called inferior races? How would we feel if Martians suddenly became the dominant species and took the same line towards us? For Wells's central characters, such speculative ideas often become matters of life and death. Trained as scientists, they not only struggle to survive the outlandish adventures Wells inflicts on them, but, while doing so, they also formulate hypotheses to account for their experiences — only for these explanations to be called into question by the

next set of fantastic events.

As his career progressed through the first half of the twentieth century, the scope of Wells's success broadened. He became internationally famous for comic novels such as *The History of Mr Polly*, social novels such as *Tono-Bungay*, and numerous books calling for a more rational society, among which the *Outline of History* stands pre-eminent. By now Wells was less interested in exploring the ethical implications of scientific knowl-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

H. G. Wells: relished the unfamiliar perspectives science opened on the human condition.

edge than in suggesting that it was the destiny of science to lead the human race to a socialist Utopia. (His 1936 film, *Things to Come*, is among the best known, and weakest, expressions of this idea.) But advisers such as Gregory and Edwin Ray Lankester ensured that Wells stayed in touch with serious scientific developments. *The World Set Free*, an inferior story published in 1914, actually foretells atomic power and the atomic bomb. In 1930 Wells produced *The Science of Life*, an outline of biology for the lay reader, in collaboration with his son G. P. Wells, a marine biologist, and T. H. Huxley's grandson, Julian. Later in the 1930s, when Wells was working on an influential

reformulation of human rights, he once again found a welcome for his ideas in the pages of *Nature*.

Michael Foot, the former leader of the British Labour Party, has surprisingly little to say about such matters. Foot admits to "scandalous ignorance" about science, but his book is skimpy in many other respects too, passing rapidly over Wells's formative years and remaining highly selective thereafter. Foot prefers to concentrate defensive fire on those areas where Wells has come in for the greatest recent criticism. He gives major coverage to Wells's sexual relationships (in fairness, there are a lot of these to cover) and to his socialist writings, dealing with both topics thoughtfully and eloquently, but playing down some awkward questions and producing a less complex portrait than we find in, say, Wells's own *Experiment in Autobiography*. Wells's personality, ideas and writings owe much of their interest to his readiness to embrace change and to to challenge expectations. From the best of motives, Foot tends to smooth out the inconsistencies and give priority to views that resemble his own. We hear much of Wells's hatred of social inequality, for example, but nothing of his sometimes negative attitude towards the working classes or his advocacy of proportional representation for parliamentary elections. Foot tells us that the subversive speeches of Chaffery in the novel *Love and Mr Lewisham* voice Wells's criticisms of "Victorian capitalism". In fact Chaffery is equally scathing about socialism and science.

Despite this homogenization, Foot delivers a sensitive account of Wells's tangled love-life and also shows beyond doubt that he was not the racist advocate of eugenics that some have alleged. When Foot was a young journalist in the 1940s, he had many meetings with Wells. This personal knowledge helps him to portray his subject as a living human being, not a mere figure from a history book.

Considered as a testimonial to Wells by a fellow author and political activist, *HG* has many merits. As a reliable biography, sadly, it disqualifies itself by vagueness, omissions and occasional errors. Chapter 2 contains two wrong dates and, among other deficiencies, fails to explain in adequate detail how Wells escaped menial work to become a science student. There is a remark about "friendly angels" that would make sense only if it referred back to an earlier discussion of *The Wonderful Visit* — one of several signs of hasty completion and poor editing. Unclear citations and the absence of a bibliography compound the problems. Committed

Wellsians will find the book repays their attention; other readers would do better to consult Norman and Jeanne MacKenzie's comprehensive study *The Time Traveller* (US title, *H. G. Wells: A Biography*). □

Michael Sherborne is Head of English at Luton Sixth Form College, Luton, Bedfordshire LU2 7EW, UK, and under his previous name, Michael Draper, is the author of *Modern Novelists: H. G. Wells* (Macmillan/St Martin's Press, 1987).

■ A more critical plea for the importance of Wells's literary contributions has also recently been published, but unfortunately did not arrive in time for inclusion in this review. In *Shadows of the Future: H. G. Wells, Science Fiction and Prophecy* (Liverpool University Press, £25 (hbk), £14.95 (pbk)), Patrick Parrinder stakes out claims for Wells as a prophet novelist, seeing Wells's legacy as the whole corpus of twentieth-century British and American science fiction.

What it's all about

Christopher Longuet-Higgins

The Artful Universe. By John D. Barrow. Oxford University Press: 1995. Pp. 274. £18.99, \$27.50.

A FEW years ago the United Kingdom was captivated by a science-fiction saga entitled "The Hitch-Hiker's Guide to the Galaxy". The hero finds himself hurtling through space in search of the secret of Life, the Universe and Everything. The secret turns out to be the number 42. In

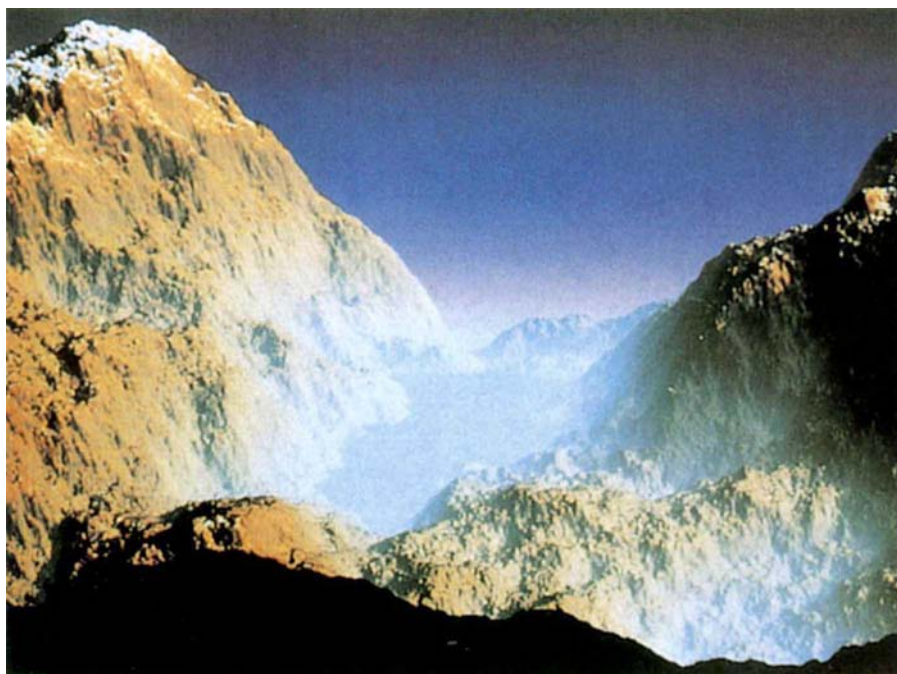
his new book John Barrow embarks on a similar quest, but with a less definitive outcome. Unlike his fellow cosmologist Roger Penrose, author of *The Emperor's New Mind*, he is sufficiently artful to cut a revolutionary figure without saying anything remotely controversial. Barrow also has the wordly wisdom to eschew mathematical equations in favour of illustrations and colour plates, essential for a grand tour that starts with the role of perspective in the visual arts and proceeds, by way of Darwin's theory of evolution, to the anthropic principle, taking in environmental concerns on the way; and that is only the first leg of the journey. (The anthropic principle — the subject of one of Barrow's earlier books — asserts that the Universe must be much as it is, otherwise we would not be here to note the fact.)

At first the argument seems to be leading somewhere — to be exploring the relationship between aesthetics and physics. Barrow makes the point that it was in the arts, not the sciences, that the role of the observer first rose to prominence. But if Renaissance painting exploited the principles of perspective, Chinese art has always favoured a more indefinite viewpoint. Barrow suggests an analogy with the subjective-objective opposition that has played such a central role in the recent history of physics. But his enquiry soon appears to be much wider in scope. A quick survey of Kant's theory of spatial perception and the non-Euclidean geometries of the great nineteenth-century mathematicians leads unexpectedly into an account of the opposing biological theories of Lamarck and Darwin. The three standard preconditions for Darwinian evolution, namely

variation, inheritance and selection, are paraphrased by three "requirements" carefully tailored to cover social and cultural as well as genetic evolution. But has human evolution in the Darwinian sense been brought to an end by cultural transmission? On such matters, Barrow warns, opinion is divided.

There follows a discussion of language and its origins, which are safely wrapped in prehistoric obscurity. After comparing and contrasting the ideas of Jean Piaget and Noam Chomsky, Barrow ventures the suggestion that "Chomsky appears to have an ambiguous attitude towards the origin of our universal grammar". But before there is time to contemplate this embarrassing lapse we are whisked in rapid succession through a sequence of essays on "the evolution of mental pictures", "the care and maintenance of a small planet" (Barrow has a jackdaw's eye for a glittering phrase), "gravity's rainbow" (his own stamping ground), "death and immortality" and "the human factor". The take-home message of this bewildering chapter is presumably the one spelt out in the introduction: "If life is to be possible within them, universes must have particular forms. When conscious life does emerge, its experiences and conceptions are strangely influenced by the fact that the Universe must be big and old, dark and cold." (Who said that scientists were not poets at heart?)

As the acceleration builds up, it becomes more and more difficult to discern a coherent train of thought, and one is forced to the conclusion that the title of *The Artful Universe* is no more than an ingenious excuse for publishing in a single book Barrow's assorted reflections on a variety of eminently fashionable but essentially unconnected topics. But perhaps it doesn't really matter whether there are any general principles — such as those of complexity theory, that elusive discipline — uniting the size of animal populations with the aesthetics of computer art (two of the main subheadings in the chapter on "size, life, and landscape"). One can skip the general waffle and turn with pleasure to Barrow's observations (he is a professor of astronomy) about the Sun, the Moon and the stars, and about the intricate connections between the week, the month, the seasons, the traditional constellations and the signs of the zodiac, not to mention the religious observances associated with these and other celestial phenomena. In the last main chapter Barrow turns to music, with the questions: "Why do we like it?" "Where does it come from?" and "Full of sound and fury, does it signify anything?" In the course of not answering these tricky questions (although he is quite kind about this reviewer's ideas on tonality) the author suggests that the appeal of music may



From *The Artful Universe*

Jagged edge: Barrow argues that our affinity for computer-generated fractal landscapes has been honed by natural selection (image by Richard Voss).

have something to do with its 1/f power spectrum; but the best part of the chapter is about something entirely different, namely the competing theories of mathematics — formalism, inventionism, Platonism and intuitionism. Possibly expecting an open verdict on *The Artful Universe*, Barrow disarms criticism at the outset by quoting the words of the philosophical taxi-driver who picked up Bertrand Russell one evening. "I said to him: 'Well, Lord Russell, what's it all about?' and, do you know, he couldn't tell me". □

Christopher Longuet-Higgins is at the Centre for Research on Perception and Cognition, University of Sussex, Brighton BN1 9QG, UK.

Designer ecosystems

Stuart L. Pimm

Miracle Under the Oaks: The Revival of Nature in America. By William K. Stevens. Pocket: 1995. Pp. 333. \$22.

ECOLOGISTS revere Douglas Adams's character Slartibartfast because designing planets is the grandest experiment imaginable in any future lifetime. Even in the present, recreating ecosystems seems more the stuff of fiction than practice. Such experiments span decades and thousands of hectares and species. William Stevens's heroes are mostly amateurs who have yet to learn that ecological experiments involve no more than 3 species, 3 years and 100 m² if they are to be published.

Illinois is an unlikely venue for exciting ecological experiments. Natural ecosystems cover less than 0.1 per cent of the state. North Branch is a strip of land running 20 km northwards from Chicago, protected from building early in the century, then abandoned to weeds. Chicago's suburbs are on my shortlist for ecological hell,

the last place to look for another ecologist. Nor would I have recognized one in Steve Packard as he stepped off his bicycle in 1975 to examine the scene. Packard's credentials were a degree in social psychology, a decade of organizing grass-roots anti-war protests and a new-found passion for wildflowers. Exactly the right skills, of course, to persuade a mix of factory workers and professionals to spend their Sundays clearing out European buckthorn and abandoned cars and planting seeds.

The North Branchers combed the snipets of surviving prairies, along railways and in cemeteries, bagging the seeds and over-wintering them in unheated garages. August 1978 was a time for assessment. Where they had cleared dogwoods, bare soil remained. Animals had eaten the small, transplanted plants. "A disaster", Packard recorded. Without fire, it would remain so.

Fire posed several problems. Safely torching areas close to homes was just the most technical. Some Chicagoans cherished alien trees. Prairies were the stuff of myth, not memory. From scientists, predictably, the advice was for more research. Practically, neither the time nor the money was available; the North Branchers put their restorations to the flame. The results were immediate and dramatic. Prairie plants flourished. Weeds retreated except under the native oak trees. Here the ground grew alien thistles and dandelions. This was all the more galling, because the once extensive 'barrens' — prairies with scattered trees — were described historically as visually striking and ecologically special.

Packard faced both political and ecological problems. By 1984, he was working for the Nature Conservancy, a nonprofit organization that has protected more than two million hectares of America's last great places by buying them and, when possible, re-selling them to government agencies. It was not then known for the intensive restoration of America's least great wild places.

The ecological problem was to find the right species mix for the barrens. Critics recommended surviving prairie remnants as sufficient models. Packard rejected these as too degraded. Early this century, Arthur Vestal had speculated that barrens lacked some characteristic prairie

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Spring clean: shooting star and white oak flourish in the spring after clearing and burning.

species and had their own distinctive species. To identify such plants, Packard looked at habitat descriptions in the *Illinois Flora*. Many would now be endangered. The crucial discovery was a country doctor, S. B. Mead, who published a list of barrens' plants in 1846. The scientific names had changed in 150 years, but by translating them from flora to flora Packard found many of his putative barrens' species in Mead's list.

Packard continued with his research-through-restoration agenda. By 1991, he had restored 136 native plant species, including whole patches of locally endangered small sundrops. Aldo Leopold coined the first law of ecological restoration. His admonition to save all the pieces when tinkering urges us to "keep all the species". Mead's list illustrates the second law: "don't lose the assembly instructions!"

Stevens covers many other restorations; most involve prairies. Bob Betz's re-creation of prairies inside the accelerator ring at Fermi Laboratory is one spectacular example. Saving prairie species has to involve restoring ecological communities and processes to areas too large and too remote from natural remnants to recover on their own. Prairies raise general problems. Is there a 'right' community to restore? In Illinois, the pre-Columbian landscape was human-dominated with fires set by the Potawatomis. Should the target communities be ones that depend on fire? The answer is not entirely arbitrary. Wrong choices drive native species to extinction. Without the right species mix, restorers must constantly weed and re-seed.

Importantly, the book is not just about restoration. Is re-assembly of complex ecosystems (or, generally, any complex system) possible? Play the Holocene tape again and will we find the same species in the same places? Is there more than one target community to restore? Indeed, might the only way to save a full complement of species be to save a diversity of communities, including ones with no historical precedents? And just what use can be the usual ecological experiments when complete replication would exceed the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fire works: the results of burning were immediate and dramatic.

space available to us and their duration exceeds the time endangered species can persist without our intervention?

Stevens is well known as a writer for the *New York Times's* weekly science section who covers a wide range of issues with rare insight and clarity. Why pick restoration as the subject for a first book? The practical difficulties, the political battles and, as the cover puts it, the "story of perseverance and tenacious human spirit" make a great read, of course. The subject matter lends itself to delightful illustrations. Crucially, the ecological issues are ones we cannot ignore. □

Stuart L. Pimm is in the Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996, USA.

Flaming nuisance?

Peter D. Moore

The Ecology of Fire. By Robert J. Whelan. Cambridge University Press: 1995. Pp. 346. £42.50, \$79.95 (hbk); £15.95, \$27.95 (pbk).

THERE can be no doubt that fire is catastrophic, but can it also be constructive? As children we are taught to avoid it, but as ecologists we need to learn how to understand it, how to control it and perhaps even how to use it; such is the message of this book.

Fire has undoubtedly played an important part in the evolution of life on Earth, long before our species began using it for social, gastronomic and hunting purposes. The traces of burned plant fragments in sedimentary rocks dating back to the Devonian period, some 360–410 million years ago, have even been used to trace the oxygen content of the Earth's atmosphere. Both the individual species and the communities that today occupy the world's fire-prone areas have therefore evolved and assembled themselves under its influence and selective direction. From the thick insulating bark of *Eucalyptus* trees to the burrowing proclivities of chaparral mammals and the evasive behavioural tactics of Australian woylies, the range of techniques that species have evolved and adopted to tolerate fire have been set out in this account. The 'Bambi' image of stampeding hordes of vertebrates fleeing before fire is not in favour with fire scientists, for animals often use a more subtle approach and double-back through gaps to pass behind the fire-front.

The dynamics of whole populations provides a documentation of benefits of fires in many habitats. Fire following fruit maturation, for example, may release

seeds for germination that would otherwise have been eaten by insects or birds, so post-germination survivorship is often greater on burned sites. But post-fire habitats offer rich pickings to wandering herbivores, as the canny Mesolithic pyromaniac hunters of northern Europe were well aware, so new shoots in burnt-over areas may be intensively grazed. This can lead to some animal populations being favoured by fire; and having considered the responses of vegetation to periodic combustion, Robert Whelan also looks at the impact of fire on animals.

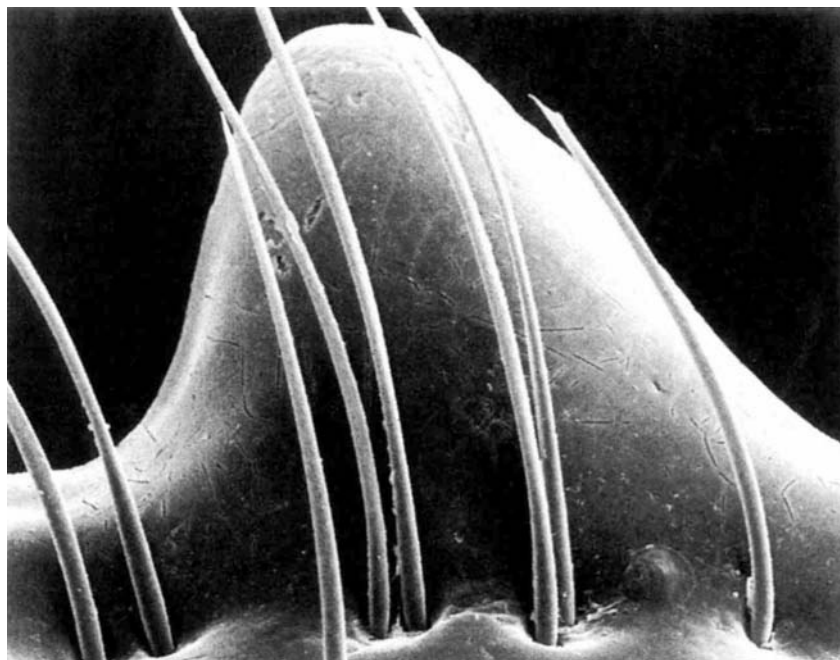
A major problem is that so much evidence about the impact of fire is anecdotal. This can be overcome in part by experimentation, which can easily be applied to plants by burning areas at different intensities and observing vegetative recovery and subsequent seed germination. But such work with animals, especially vertebrates, would quite reasonably be frowned upon. Sound information on the mortality inflicted on large herbivores by fire is therefore still scarce. Among invertebrates, results cited vary from an almost total kill in a forest fire to about 90 per cent survival by flight for savanna grasshoppers.

The book provides a thorough analysis of the ecology of fire, illustrating principles from a wide range of fire-affected habitats, including boreal forest, Mediterranean-climate chaparral

and savanna grassland. Europeans may be disappointed at the poor representation of the oceanic western heaths and moorlands in the text, and this is particularly apparent in the final chapter on fire and management. Here we find what should be the climax of the book, an explanation of how we can use this powerful tool to manipulate ecosystems for habitat management and wildlife conservation, but the attitude conveyed by the few pages devoted to this subject still seems to view fire as a danger to be avoided rather than an ally to be exploited. The main emphasis of the chapter is on hazard reduction by periodic burning in American forests, and this is certainly an important topic that demands attention. An account of the historical development of fire policies and the disastrous mistakes made (as at Yellowstone National Park) would have been appropriate here to illustrate the dangers of over-protectionist attitudes. But we also need more on the positive side of fire management, and where better to turn than the British grouse moors with their long history of fire management? Some conversations with Scottish muir-burners would have turned this otherwise admirable text into a more rounded work. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

The art of the unseen



SENSORY hairs protruding from the head of a "patent leather beetle" (magnification, $\times 200$). The picture is one of 200 or so images that appear in *Journeys into Microspace* by Dee Berger, a skilful exponent of the art of scanning electron microscopy. Nylon stockings, pancreatic stones, dinosaur teeth and sticky tape, to name but a few, will never seem the same again. Columbia University Press, £39.50.

Role of hydration and water structure in biological and colloidal interactions

Jacob Israelachvili & Håkan Wennerström

The conventional explanation of why hydrophilic surfaces and macromolecules remain well separated in water is that they experience a monotonically repulsive hydration force owing to structuring of water molecules at the surfaces. A consideration of recent experimental and theoretical results suggests an alternative interpretation in which hydration forces are either attractive or oscillatory, and where repulsions have a totally different origin. Further experiments are needed to distinguish between these possibilities.

WATER is the natural solvent. The vast majority of molecular interactions in living systems occur in an aqueous environment, and geological structures are constantly exposed to water. Because of increasing concerns about the environment, water is gradually replacing organic solvents for technological and household products such as water-based fuels and lubricating fluids. The factors that determine why water is often, but not always, a good solvent or suspending medium for colloidal and biological interactions is the focus of this Review.

It has long been appreciated that both as a pure liquid and as a solvent, water is complex and has some unusual if not unique properties¹. The complexity of liquid water is due to a combination of the small size and distinct polar charge distribution of the water molecule². The charge distribution can be modelled by four charges located along the four arms of a tetrahedron^{2,3} which allows each water molecule to participate in strong polar (electrostatic charge-dipole or hydrogen-bonding) interactions⁴ with a high degree of spatial directionality. The strong hydrogen-bonding water–water interaction results in a large cohesive energy or latent heat, a high boiling point, a high surface tension, and a reluctance to dissolve inert (nonpolar or hydrophobic) solutes with which it cannot interact through similarly strong polar forces⁵.

But water can also strongly bind to and dissolve polar and hydrophilic compounds, and our main concern here will be to investigate those properties of water that make it such an excellent solvent for a wide variety of solute molecules and ions, and for suspending colloidal particles and biological structures such as proteins, DNA, viruses and cells. To explain the powerful solvent properties of water, scientists naturally turned to consider the intermolecular forces acting between the dissolved species in water and aqueous electrolyte (salt) solutions. The two major forces operating between two macromolecules or surfaces in liquids are the attractive van der Waals and repulsive electric double-layer forces⁶. The former is always present; the latter depends on the existence of charged surface groups. But even uncharged molecules and particles are often miscible with water—an observation that led Langmuir⁷ and Derjaguin⁸ to postulate the existence of an additional repulsive force that seemed to be unique to water. This hydration or structural force is believed to arise from the strongly bound and oriented first layer of water molecules on surfaces (Fig. 1 *top*) which may prevent two surfaces or macromolecules from approaching any closer than 5–6 Å—the thickness of two water molecules. There was good experimental and theoretical evidence for the existence of such a primary hydration shell or layer. But to explain a repulsion that exceeded the van der Waals attraction in both magnitude and range, this force had to propagate further than one or two molecular layers. It seemed natural that the first oriented layer

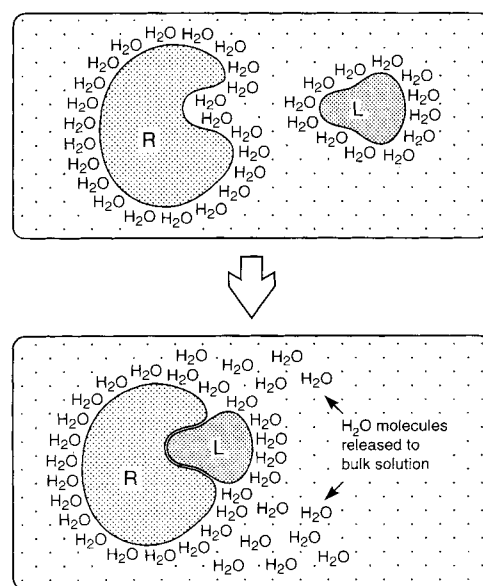


FIG. 1 Schematic illustration of a receptor molecule (R) binding to a ligand group (L) in an aqueous medium.

could induce a second layer to orient, the second would likewise influence the third, and so on. A picture emerged of hydrophilic surfaces bounded by a coat of structured water that opposed being disrupted and thus gave rise to a monotonically repulsive force between two approaching surfaces in water. The range of this interaction was variously suggested to extend from nanometres^{9–12} to many micrometres¹³, and various theories based on water structure were proposed^{14–17}. Interestingly, until the early 1980s, with reports that the hydrophobic interaction may be long-ranged^{18,19}, explaining unexpected *attractive* interactions does not appear to have been a serious problem, this is because at least one type of attractive force—the van der Waals force—is always expected to be present.

The interpretation of many phenomena in terms of an extensive protective layer of structured water around molecules and surfaces leads naturally to certain implications and consequences. Thus, to change the hydration forces in water, emphasis was placed on modifying not the structure of the hydrated surfaces, but that of the water itself, for example, by adding 'structure makers' or 'structure breakers' (such as salts) to the solution²⁰. In colloidal and biological systems, the idea that the hydration layer must be overcome before two molecules, colloidal particles or

membranes can come into contact and react, adhere or fuse is prevalent, again with the implications that factors that alter the water structure—for example, by dehydrating the surfaces—are required to initiate these phenomena⁹.

An alternative possibility is that the origin of short-range stabilizing repulsions do not arise from water structuring effects associated with some peculiar property of water–water interactions intrinsic to water, but are more to do with the chemical and physical nature of the surfaces. With this new interpretation of the origin of these forces, many colloidal and biological interactions in water are seen in a completely different light, with different mechanisms suggesting themselves for controlling phenomena such as particle adhesion, membrane recognition and fusion, biochemical reactivity and the rates of complementary associations. Here we review the experimental evidence that leads to this suggestion, and compare its explanatory power with that of the standard picture of hydration in terms of water structure. We suggest that any interaction that may arise from water structuring effects is expected to be monotonically attractive or oscillatory, not repulsive. The alternative model implies that water has more in common with other molecular liquids than is often supposed. But a firm resolution of the question of which model is to be preferred will have to await further experiments and theoretical modelling, some of which are suggested at the end of this Review.

Molecules and surfaces that associate in water

The total interaction between any two solute molecules or particle surfaces in water involves, in addition to the direct interaction, an additional interaction arising from the displacement of the solvent (water) molecules as the two suspended molecules come into contact (Fig. 1). Thus, the overall energetics of any association or dissociation process in water also involves what is commonly referred to as dehydration or hydration of the interacting surface groups. But it has never been an easy matter to establish the mechanism behind this interaction, or even whether it is the major contributor. For example, it is widely believed that the DNA double helix owes its stability to the hydrogen bonds holding the two strands together. However, in the dissociated state, the individual strands can form even stronger hydrogen bonds with the solvent water molecules²¹, so the hydrogen bonds can be only part of the picture. The overall energy balance that favours the association of two DNA strands and other macromolecular associations in water depends on a competition between solute–solute, solute–solvent and solvent–solvent bonds (including hydrogen bonds) the outcome of which is generally not obvious or simple to analyse.

Similarly, the macroscopic surfaces of certain colloidal and biological surfaces that expose hydrogen-bonding sites (such as bilayers of fatty acids²²) or a surface layer of ordered water molecules (such as brushite crystals, $\text{CaHPO}_4 \cdot 2\text{H}_2\text{O}$; refs 23,24) associate spontaneously in water, expelling any bulk liquid water into the solvent reservoir as they do so. In all of the above cases, we would say that the dehydration of the surfaces or exposed molecular groups is energetically favourable, and that it results in an overall attraction between these groups in water.

The full picture is often more complex because additional attractive nonpolar van der Waals and hydrophobic interactions between surface nonpolar groups will also contribute towards, or even dominate, the overall attraction. For example, the attraction between the nonpolar nucleotide rings in DNA is stronger than the hydrogen-bonding interaction between the polar groups^{25,26}, so that the hydrogen bonds should be seen as providing the structural specificity of the DNA double helix but not the major driving force of the association²¹.

Molecules and surfaces that are stable in water

In other cases, strongly hydrophilic molecules (certain proteins, viruses and DNA) or surface groups (such as the polar head groups of surfactant and lipid molecules) are found to effectively

repel each other in water; that is, the structures (micelles, vesicles) remain dispersed in water. In this case we would say that the dissociation or repulsion arises because their hydrogen-bonding interaction with water is stronger than the solute–solute interaction, that is, it is due to their favourable hydration interaction. But here too, additional repulsive forces may be involved that must be taken into account. These include the electric double-layer forces between any charged surface groups^{6,19}, and the entropic or thermal fluctuation forces between any mobile surface molecular groups^{27,28}. Either of these forces could dominate over the other interactions (both attractive and repulsive) and govern the overall repulsion.

The repulsive electric double-layer force between similarly charged surfaces in electrolyte solution is generally attributed to the electrostatic repulsion between the surface charges. But as has been stressed in the literature²⁹, the purely electrostatic contribution to the net interaction is actually attractive—the net repulsion arising from the entropy or osmotic pressure of the counterions between the surfaces. This important conceptual point can be explained simply as follows: the diffuse atmosphere of counterions surrounding a charged surface (Fig. 1 *top*) arises, not because of the electrostatic interaction between the surface co-ions and oppositely charged diffuse-layer counterions, which is attractive and therefore opposes dissociation, but because of the overriding effect of the increased entropy of the counterions when they leave the surface. When two such surfaces approach each other, the counterions are pushed back onto the surfaces, and work has to be done against this entropic (osmotic) pressure; but as far as the electrostatic processes are concerned, the pushing of the counterions back onto the oppositely charged surfaces is actually favourable. The entropic, rather than electrostatic, origin of the electric double-layer repulsion is significant when considered with the other types of interactions that give rise to an effective repulsion between molecular groups or surfaces in water, and we return to this matter and its implications in the final section.

Measurements of forces between surfaces

In recent years, experimental techniques for directly measuring the forces between surfaces in liquids have matured. We can now

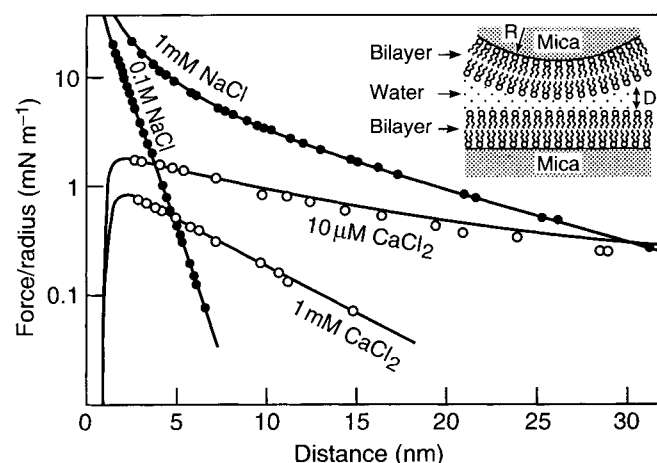


FIG. 2 Typical DLVO force profile measured between two smooth charged surfaces in aqueous electrolyte solutions. The force, F , normalized by the surface radius, R , is plotted against surface separation, D . The data shown here are for two negatively charged bilayers composed of phosphatidylglycerol lipids⁸³, which are common constituents of biological membranes. The agreement between experiment (circles) and the DLVO theory (lines) is very good, and shows that the overall interaction is dominated by repulsive electric double-layer forces at large separations and attractive van der Waals and ion-correlation forces at small separations^{3,19,84}. The maximum at B is known as the force or energy barrier. For many systems, the continuum DLVO theory has been found to describe satisfactorily the surface interactions down to separations of 1–2 nm (refs 19, 27, 30, 73).

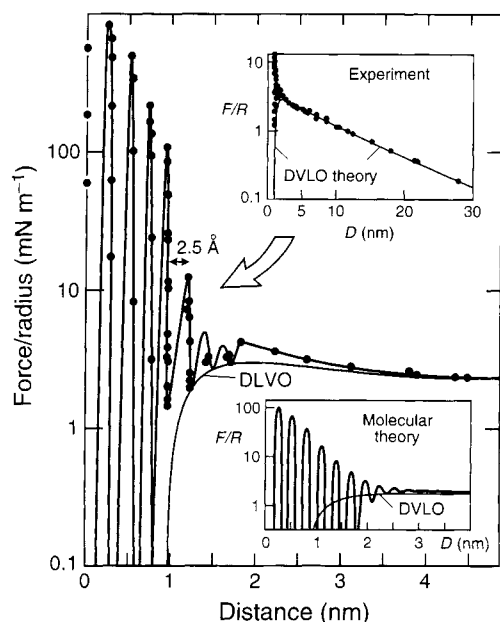


FIG. 3 Short-range deviations from continuum DLVO theory measured between mica surfaces in 1 mM KCl solutions^{36,37}. Below 20 Å (about eight water-molecule diameters) the force becomes increasingly oscillatory, with a periodicity equal to the diameter of water molecules (2.5 Å), indicative of the existence of diffuse water layers between the surfaces. Inset, theoretically predicted force for this system at small separations, based on a non-continuum molecular theory⁵¹.

accurately measure the forces between smooth solid surfaces, lipid bilayers or model biomembrane surfaces containing proteins and other biological molecules using the surface forces apparatus^{3,19,30} or osmotic pressure technique^{9,31,32} and between microscopic colloidal particle surfaces using the atomic force microscope³³. In the osmotic pressure technique^{9,31,32} inter-bilayer forces are measured by osmotically withdrawing water from a dispersed stack of bilayers in solution while measuring the change in the lamellar repeat spacing with X-rays. In the surface forces apparatus^{3,19,30}, the forces between two surfaces immersed in a liquid are measured from the deflection of a spring supporting one of the surfaces, while the distance between the surfaces is measured using an optical interference technique. As an example, Fig. 2 shows the measured forces between two charged lipid bilayers in aqueous NaCl and CaCl₂ solutions. At long range, there is a repulsion that increases exponentially with decreasing separation. At a separation of $D \approx 2.5$ nm, the repulsion turns into an attraction. This behaviour of the forces was predicted fifty years ago by Derjaguin and Landau³⁴ and Verwey and Overbeek³⁵, and the so-called DLVO theory has since formed the basis for understanding colloidal interactions in liquids, especially aqueous solutions. The salient features of this theory are that the long-range repulsion is due to the electric double-layer force, whereas at shorter range the attractive van der Waals force wins out.

When we consider the approach of two real surfaces or macromolecules that interact with an equilibrium interaction potential as shown in Fig. 2 (the interacting species could be two clay particles, a ligand–receptor pair or a DNA–repressor pair), it is clear that the dynamics of the association process will be strongly influenced by how the interaction energy or force varies with distance. Thus, even if it is energetically favourable for the two species to come into contact (high binding energy at $D = 0$), the association may be kinetically inhibited due to the existence of the repulsive force or energy barrier at some finite separation (compare Fig. 2). This can give rise to slow aggregation even when the

bound state is highly favourable. Conversely, an association process of low binding energy can be anomalously rapid if there is a long-range attractive force that steers the two partners together.

As the experimental techniques have improved and as more systems are studied, it has become evident that, although very useful as the reference interaction, the DLVO theory leaves many observations unexplained^{19,30}. Figure 3 shows the measured force between two mica surfaces in aqueous KCl solutions^{36,37}. The long-range force, beyond about 2 nm, is well described by the DLVO theory, but careful measurements reveal oscillations in the force at short range. The period of these oscillations is ~ 0.25 nm which corresponds to the size of a water molecule. The water molecules appear to be partially arranged in layers, and as two surfaces approach these layers are squeezed out sequentially. Such short-range ordering occurs for all simple liquids between smooth, rigid surfaces and it is not exceptional for water or smooth mica surfaces¹⁹. The oscillatory force is always expected to be attractive at contact ($D = 0$)^{38,39}.

Water structure effects

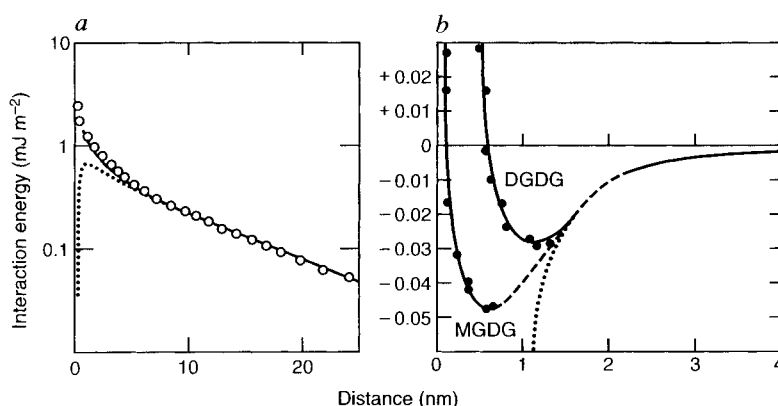
Theoretically, for two smooth surfaces interacting in an aqueous solution by a combination of DLVO and oscillatory forces, the contact interaction in the limit of $D \rightarrow 0$ is always expected to be attractive^{38,39}. However, the most detailed measurements of the forces between some surfaces, notably between silica surfaces and uncharged lipid bilayers, show a monotonically increasing repulsive force in this range (Fig. 4). What is the source of this repulsion?

When confronted with such experimental findings investigators have taken refuge in the old idea of water structure, envisaging a layer of structured water molecules, extending some 1–3 nm or more from each surface, which provides a protective force barrier against the approach of two ‘hydrophilic’ surfaces or groups. This interpretation seems reasonable for a number of reasons. Water is known to have strong and directional intermolecular interactions and an unusually high degree of structure in the liquid phase at ambient temperatures. It thus appears conceivable that the extra monotonically repulsive force between two particles in water could arise from a particular structure or order that is propagated through the solvent. Through the years this idea has had a remarkable appeal. When confronted with unexpected experimental results water structure has commonly been used as a *deus ex machina* for explaining the observations^{40–43}. The temptation is not new: both long-range and short-range forces arising from water structure were controversial issues already in the 1920s and 1930s (compare Porter forces⁴⁴) as the following quotation from Hartley’s 1936 classical book⁴⁵ on surfactants illustrates: “There is a widespread tendency to use ‘hydration’ in colloid chemistry as a sort of universal explanation of puzzling phenomena. Its inaccessibility to direct experimental determination fortifies this tendency.”

The attribution of all short-range repulsions in water to surface hydration layers has fostered a whole family of concepts that by now have become well established—that other properties of water such as its diffusivity and dielectric constant must be very different near a surface; that other hydrophilic groups cannot contact each other or chemically interact when protected by a hydration layer, and that before certain adhesion and fusion events can occur, two colloidal particles or biological membranes must somehow overcome their hydration force barrier. The search for salting-out and dehydrating agents, such as water structure breakers, has followed naturally from this hypothesis, as does the current search for a satisfactory theoretical description of water structure at surfaces.

Marčelja and co-workers^{14,15} made an ingenious attempt to construct a detailed molecular model of water-mediated structural interactions that accounted for some of the experimental results of the type shown in Fig. 4. The Marčelja theory, formulated in terms of a decaying water structure from each surface,

FIG. 4 Measured forces F (expressed as interaction energies, $E = F/2\pi R$) between surfaces in aqueous solutions that appear to show an additional monotonic repulsion at small separations, below ~ 20 Å, where the DLVO theory would predict an attraction down to contact (at $D = 0$). *a*, Two silica surfaces^{62,68}. Dotted line, DLVO theory; solid line, modified DLVO interaction for the case where the planes of origin of the double-layer and van der Waals forces are shifted by 5 Å relative to each other⁶², as illustrated in Fig. 5*a*. *b*, Two supported uncharged glycolipid bilayers of monogalactosyl diglyceride (MGDG) and digalactosyl diglyceride (DGDG) in water⁸⁵. At distances beyond the adhesion minima the measured forces are well described by the DLVO theory, which consists of a pure van der Waals attraction in this case of two uncharged surfaces. Similar monotonic repulsions have been measured between other lipid bilayers^{27,54}.



predicts an exponentially decaying repulsive force of a magnitude that is determined by the degree of orientation or polarization of the first layer of water molecules, and with a decay length that is a characteristic property of the liquid (water). This theory is widely quoted whenever a short-range monotonic repulsion is found in water. However, it has been increasingly clear that the Marčelja model is unsatisfactory: computer simulations do not support the structural effects^{28,46–48}; measured forces are too system-dependent in both magnitude and distance-dependence (often not exponential)^{27,49}; fitted decay lengths in water vary by more than an order of magnitude^{49–53}, from 0.6 Å (ref. 50) to above 6 Å (refs 49,53); the absence of an adhesive minimum at contact is unexplained; the observed temperature effects^{27,32,51,53–55} remain a mystery; bilayer surfaces have been found to be highly diffuse and disordered^{48,56} and recent measurements show that these forces are not specific to water^{57,58}. It is also fair to say that no theory, based on water or solvent structure, has been proposed that is predictive, theoretically sound and that accounts qualitatively for the observed effects.

Effects of surfaces on interactions in liquids

Recent experiments suggest that to understand how repulsive short-range forces of the type shown in Fig. 4 arise, it is necessary to focus on the detailed structure and properties of the interacting surfaces or interfaces rather than on the structure of the solvent between them. Although the forces in Fig. 4*a* and *b* show some similarities, careful measurements have revealed that they may have different molecular causes. Silica is a solid with uncharged silanol (Si–OH) and charged silicic acid (Si–O⁻) groups at the surface⁵⁹. Protons can dissociate from the surface silanol groups, generating a DLVO-type repulsive double-layer force at long range. Furthermore, the silica surface is generally amorphous and not molecularly smooth: a low density of molecularly thin silica ‘hairs’ stick out from them⁶⁰. This pushes the double-layer repulsion farther out relative to the van der Waals attraction at the same time as a short-range steric repulsion is generated due to the protruding hairs (Fig. 5*a*). This mechanism, first proposed by Frens and Overbeek⁶¹, can explain quantitatively the observed monotonic short-range repulsion and absence of adhesion in the case of silica⁶². Such protruding surface groups should not be considered as fixed or rigid, but more as a surface layer of flexible, polymer-like segments, that can be expanded or collapsed by changing the solution conditions, as occurs with adsorbed polymers⁶³. For example, in the case of the negatively charged ‘co-ion’ groups on silica, these would be expected to collapse in the presence of multivalent metal ions such as calcium⁶⁴ thereby also reducing the short-range repulsion, as observed⁶⁵.

Similar mechanisms, but in terms of the finite sizes of the adsorbed hydrated counterions, rather than the co-ions, have been shown to account quantitatively for the enhanced monotonic repulsions measured between mica surfaces in concentrated electrolyte solutions^{66,67}. Again, an outward shift of the surface

charges by only a few ångströms is sufficient to remove the van der Waals adhesion at contact, and replace it with a strong short-range repulsion that appears to extend some tens of ångströms; and the finite sizes of the hydrated counterions further enhance the repulsive osmotic component of the double-layer force via an excluded volume effect⁶⁷. In all of the above cases, the experimental results could be accounted for quantitatively without invoking any additional repulsive hydration force. We do not consider the first layer of water molecules bound to a surface as producing a hydration force or solvation force. This primary hydration shell has nothing to do with water–water but rather with water–surface interactions, and is therefore not a characteristic or intrinsic property of water. Such solute–solvent interactions arise in all non-ideal binary systems resulting in excluded volume effects that are well understood physically and thermodynamically. Recent direct measurements of the viscosity of water and electrolyte solutions adjacent to silica⁶⁸ and mica⁶⁹ surfaces have also shown that the non-slip plane is located no further than one water layer from these surfaces—again consistent with the absence of water structuring at these surfaces.

In the case of surfactant and lipid bilayers, above the chain melting transition the molecules have significant thermal flexibility within the bilayer. At the interface molecules protrude out from the bilayer^{48,56,70} head groups change conformation²⁸, and there are also collective undulation modes^{71,72} (Fig. 5*b*). In the presence of an opposing surface these motions become restricted, resulting in a monotonically repulsive force^{27,28,48,72}. The similarity in the range and magnitude of the short-range repulsive forces measured between certain bilayers and various other macromolecules and surfaces across water has often^{9,31} been taken as an indication that they must be due to water structure. But what was missing from this analysis, and most previous analyses, was an appreciation that the short-range entropic forces between mobile surface groups and ions are also not very different (known as the osmotic pressure limit) and that even the simplest theoretical models of entropic thermal fluctuation interactions can account, both quantitatively and qualitatively, for most of the experimental observations on amphiphilic systems^{27,28,48,73} and—in the case of lipid bilayers—down to the last one or two water molecules from each bilayer surface^{32,55}.

Forces arising from water structure

How, then, does water structure influence the forces between molecules or surfaces? As illustrated in Fig. 3, the presence of structure in the liquid generally gives rise to an oscillatory force, but only between smooth rigid surfaces. This is the basic structural force in all liquids^{19,38,74,75}. In more complex situations involving rough or chemically heterogeneous surfaces, or soft amphiphilic or biological surfaces, we should expect a monotonically varying component due to the ‘smearing out’ of the oscillations. However, experiments with other surfaces and liquids show that surface roughness and surface fluidity reduce the range and magnitude of

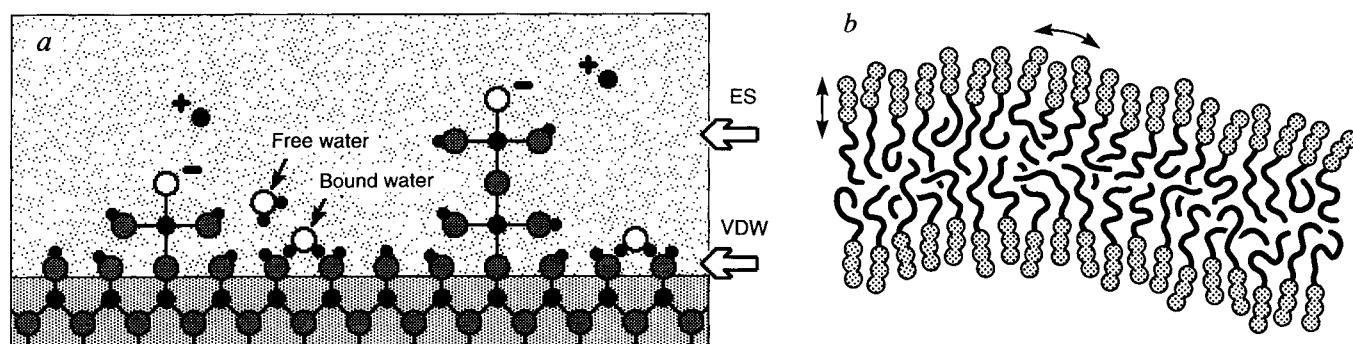


FIG. 5 *a*, Silica-like surfaces in water exposing silanol ($\text{Si}-\text{OH}$) and silicic acid ($\text{Si}-\text{O}^-$) groups, where the surface charges are located a few ångströms out from the solid-liquid interface (grey spheres, Si; white spheres, O). A difference of as little as 5 Å between the plane of electrostatic (ES) charge and van der Waals (VDW) plane is enough to totally alter the DLVO interaction between two surfaces over the last 50 Å before contact, transforming the expected short-range attraction (dotted curve in Fig. 4*a*) into a monotonic repulsion (solid curve in Fig. 4*a*).^{61,62} *b*, Lipid bilayer or

biological membrane in the fluid state in water, where the constituent molecules are in continuous thermal motion, including protrusion and headgroup tilting motions (indicated by vertical and horizontal arrows), and where the membrane as a whole also undergoes gross bending and thickness fluctuations. When two such membranes approach each other, the confinement of these thermal fluctuation modes gives rise to a short-range monotonic repulsion that is analogous to the osmotic pressure of a solution or to the normal pressure of a gas²⁷.

the oscillations^{76,77} and that in binary liquid mixtures the range is less than in either of the pure liquids⁷⁸. We have previously argued²⁷ that for amphiphilic surfaces there is no firm evidence, experimental or theoretical, in favour of a monotonically repulsive structural force. We now suggest that this may be generally true for all surfaces. First, we illustrate in Fig. 6 how a force arising from surface-induced water structuring effects could be expected to be attractive. This may be the origin of the hydrophobic interaction⁷⁹, although the experimental situation is still unclear and we do not wish to discuss the origin of this still elusive interaction here. Attractive hydrophobic forces are sometimes thought to be a negative hydration force. However, just as in the case of DLVO theory or the Lennard-Jones potential, attractive and repulsive forces can have very different origins, so the repulsive and attractive non-DLVO interactions seen in water should not be assumed to be related *a priori*.

The molecular mechanism giving rise to a structural force involving oriented dipolar water molecules is shown in Fig. 6. We consider a surface that interacts strongly with the first layer of water molecules in direct contact with it. This strong interaction may be due to physical or chemical binding, or to a good complementary fit between the water and surface groups. But regardless of the specific mechanism responsible for ordering or orienting the first layer of water molecules, this interaction must be recognized as arising from water-substrate or solvent-solute forces and not from a pure hydration (water-water or solvent-solvent) interaction. These come in only beyond the first or 'primary' hydration layer. Figure 6*a* shows the electrical field lines arising from the first layer of oriented water molecules. We may note that the field is highly non-uniform in the x, y (in-plane) and z (out-of-plane) directions. The field changes direction across one unit cell on the surface, which implies that the field cannot be described in terms of a mean-field at any point z from the surface (by applying Gauss' law, it can be shown that the average field in the z -direction is zero).

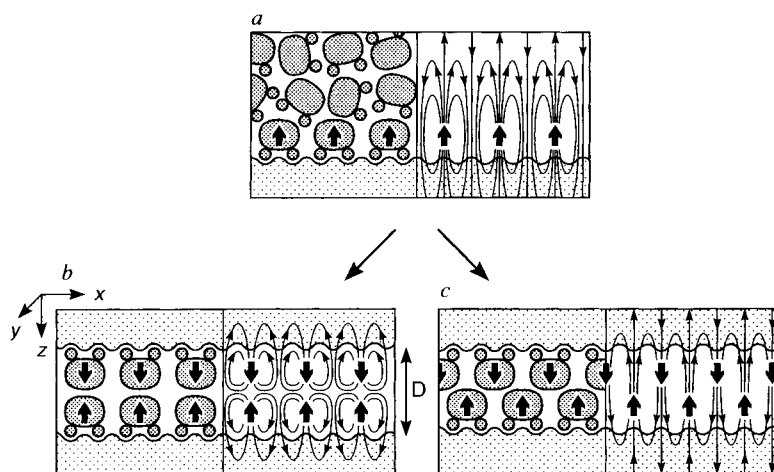
If a second surface approaches the first with its surface water molecules in perfect juxtaposition, a repulsion results (Fig. 6*b*). Thus, any theoretical solution to this problem that assumes mirror symmetry will predict a repulsion, but it will not be computing the interaction of lowest energy. When one of the two surfaces is translated a distance of half a unit cell in the x and y directions, the electric field lines from the two surfaces now add constructively resulting in an attractive interaction⁸⁰ (Fig. 6*c*). At first sight, it seems counter-intuitive that two surfaces having their water dipoles pointing towards each other will attract. The answer to this apparent paradox is to be found in the long-ranged character

of dipolar interactions. The nearest oppositely oriented water molecules do indeed repel, but the effect is more than compensated for by the water dipoles that are farther away. The complexity of this situation, where a small shift in the relative positions of the two surfaces changes a repulsion into an attraction, illustrates the subtlety that can arise when electrostatic dipolar and molecular structural effects are involved together in determining the final net interaction. Theories that simplify the system by assuming a mean-field or ignoring the possibility of asymmetric or staggered molecular ordering (that may also be continually changing as two surfaces approach each other) may miss the lowest-energy configurations at any given surface separation, and thus predict an erroneous repulsive interaction. Computer simulations, however, are less likely to fall into this trap and, for rigid solid surfaces, have so far have predicted only oscillatory or attractive forces arising from surface-induced water-structuring effects^{46,47,81,82}. Recent simulations of fluid bilayer-water surfaces⁴⁸ have also found that the decaying polarization and structure of water molecules from bilayer surfaces are determined by the surface electric fields and 'dynamic roughness' of the bilayer surfaces, with no indication of any water-intrinsic decay length.

Repulsive forces due to entropic interactions

Thus we propose that the steep short-range repulsions that are often observed between colloidal particles and biological surfaces in water are not due to a layer of structured water but to the entropic repulsion arising from the confinement of thermally mobile surface groups—the properties of water being essentially the same as in the bulk beyond the first layer of surface-adsorbed water (which may or may not be strongly bound to the surfaces). Thus, water should not be considered to have a different mobility, or dielectric constant or freezing point, near or between two surfaces (apart from effects arising from the normal freezing-point depression of solutions). Nor should one consider that there are thick hydration layers preventing proteins, ligands, receptors and other surface hydrophilic groups from coming into contact with each other. Indeed, it is this very contact that gives rise to the repulsion, which is thus not unlike the repulsive pressure generated by gas molecules colliding with a wall²⁷ (it is not always appreciated that large osmotic pressures can be produced by very low solute concentrations). In the case of colloidal, micellar, vesicle and cell-cell interactions, including membrane fusion, there is no thick hydration barrier that must be overcome before the surfaces or molecular groups on the opposing surfaces can interact directly with each other; the only real force barrier is the first layer of water molecules directly in contact with the surfaces.

FIG. 6 a, Hydrophilic surface with a strongly bound first layer of oriented water molecules, shown by the darkly shaded figures that resemble H_2O molecules. The thick arrows indicate the direction of the oriented water dipoles, and the thin arrows indicate the electrostatic field lines (shown only in the right half of each panel for clarity). It is important to appreciate that these primary binding forces, which can arise from specific ionic, hydrogen-bonding or complementary-fit interactions, act between the surface and water, and that they occur for all substrate-solvent or solute-solvent systems. These forces should be distinguished from those acting between the water molecules themselves (the solvent-solvent and electrostatic forces) which occur beyond the first layer. b, Two surfaces as in a interacting in registry in water. The resulting electrostatic interaction is an exponentially decaying repulsive force with a decay length λ equal to some fraction of the spacing d between adjacent dipoles on the surfaces ($\lambda \approx d/2\pi$ for a square surface lattice, $\lambda \approx \sqrt{3}d/4\pi$ for a hexagonal lattice)^{19,80}. This 'symmetric' situation, in which the mid-plane is a mirror plane of symmetry, is often implicitly assumed in computer simulations. c, Most favourable interaction-energy configuration, attained when the two surfaces are allowed to fully relax at any given separation. This asymmetric or 'staggered' configuration gives rise to an attractive interaction of the same magnitude and range, but of opposite sign, as in b. This lowest-energy configuration can easily be missed in simplified models of short-range electrostatic forces between



surfaces. This schematic diagram shows that the electrostatic dipole-dipole interactions alone cannot give rise to a repulsion. Inclusion of the entropy of rotating and translating dipoles reduces the strength of the attraction¹⁹ but, unlike the entropy of ions in a double layer, it does not change the sign of the interaction.

Likewise, the action of a dehydrating agent or water-structure breaker may be more on the nearby surface than on the water. More generally, one should perhaps look more towards how surfaces rather than solvents may be modified to bring about some desired effect, such as the coagulation of a colloidal dispersion or a better lubricant. And finally, as a solvent and suspending medium, water should be seen as an ordinary liquid, whose molecules have a small size and strong intermolecular interactions with surfaces and solute molecules.

Further experiments and theoretical modelling are required to test the validity and generality of the above hypothesis. Experimentally, this will require more discriminating molecular probes of surface structure, such as specially modified atomic force microscope probes that can function reliably under water, to measure the static and dynamic roughness of surfaces and check for the existence of small protruding surface groups. On the conceptual and theoretical levels, we need an improved description of the molecular-scale properties of surfaces. Theoretical treatments of interactions between surfaces at short range require that the surfaces be treated, not just as hard or soft walls, but with the same molecular detail as are the intervening liquid molecules, including a proper balancing of the interplay between the long-range and short-range intermolecular forces. Given the complexity of these systems, any quantitative analysis will have to rely on computer simulations, which have now matured so that it is becoming feasible to evaluate explicitly the forces between surfaces in addition to the molecular structure. But care must be taken to avoid artificial restrictions, such as assuming mirror symmetry of the surfaces, which can produce results that do not reflect the interaction of lowest energy (compare Fig. 6b and c). Experimental and theoretical methods have now reached a level of sophistication that the 60-year-old issue of hydration forces—so central to a molecular-level understanding of biological and inorganic systems in water—may soon be resolved. □

J. Israelachvili is at the Department of Chemical Engineering and the Materials Department, University of California, Santa Barbara, California 93106, USA. H. Wennerström is at Physical Chemistry 1, Chemical Center, University of Lund, PO Box 124, S-22100 Lund, Sweden.

1. Franks, F. in *Water: A Comprehensive Treatise* Vols 1–7 (ed. Franks, F.) (Plenum, New York, 1972–82).
2. Allen, M. P. & Tildesley, D. J. *Computer Simulations of Liquids* (Clarendon, Oxford, 1987).
3. Marčelja, S. in *Liquids at Interfaces* Ch. 2 (eds. Charvolin, J., Joanny, J. F. & Zinn-Justin, J.) 99–151 (Elsevier Science, Amsterdam, 1990).
4. Coulson, C. A. *Valence* 2nd edn Ch. XIII (Oxford Univ. Press, London, 1961).
5. Tanford, C. *The Hydrophobic Effect* (Wiley, New York, 1980).
6. Hunter, R. J. *Foundations of Colloid Science* Vol. 1 (Clarendon, Oxford, 1989).
7. Langmuir, I. *J. chem. Phys.* **6**, 873–896 (1938).
8. Derjaguin, B. V. & Churaev, N. V. in *Fluid Interfacial Phenomena* Ch. 15 (ed. Croxton, C. A.) 663–738 (Wiley, Chichester, 1986).
9. Rand, R. P. & Parsegian, V. A. *Biochim. biophys. Acta* **988**, 351–376 (1989).
10. Pashley, R. M. *Adv. Colloid Interface Sci.* **16**, 57–62 (1982).
11. Israelachvili, J. N. & Pashley, R. M. in *Biophysics of Water* (eds Franks, F. & Mathias, S.) 183–194 (Wiley, Chichester, 1982).
12. Israelachvili, J. N. *Chemica Scripta* **25**, 7–14 (1985).
13. Drost-Hansen, W. in *Biophysics of Water* 163–169 (eds Franks, F. & Mathias, S.) 163–169 (Wiley, Chichester, 1982).
14. Marčelja, S. & Radic, N. *Chem. Phys. Lett.* **42**, 129–130 (1976).
15. Gruen, D. W. R. & Marčelja, S. *J. chem. Soc., Faraday Trans. II* **79**, 225–242 (1983).
16. Schiby, D. & Ruckenstein, E. *Chem. Phys. Lett.* **95**, 435–438 (1983).
17. Attard, P. & Batchelor, M. T. *Chem. Phys. Lett.* **149**, 206–211 (1988).
18. Israelachvili, J. N. & Pashley, R. M. *Nature* **300**, 341–342 (1982).
19. Israelachvili, J. N. *Intermolecular and Surface Forces* 2nd edn (Academic, New York, 1991).
20. Drost-Hansen, W. & Clegg, J. *Cell-Associated Water* (Academic, New York, 1979).
21. Voet, D. & Voet, J. G. *Biochemistry* 811 (Wiley, New York, 1990).
22. Tiddy, G. J. T. *Phys. Rep.* **57**, 1–46 (1980).
23. Hanein, D., Geiger, B. & Addadi, L. *Langmuir* **9**, 1058–1065 (1993).
24. Curry, N. A. & Jones, D. W. *J. chem. Soc.* **A**, 3725–3729 (1971).
25. Cantor, C. R. & Schimmel, P. R. *Biophysical Chemistry* Ch. 23 (Freeman, San Francisco, 1980).
26. Smith, C. A. & Wood, E. J. *Biological Molecules* Ch. 8 (Chapman & Hall, London, 1991).
27. Israelachvili, J. N. & Wennerström, H. *J. phys. Chem.* **96**, 520–531 (1992).
28. Granfeldt, M. & Miklavic, S. *J. phys. Chem.* **95**, 6351–6360 (1991).
29. Jönsson, B., Wennerström, H. & Halle, B. *J. phys. Chem.* **84**, 2179–2185 (1980).
30. Christenson, H. K. *J. Dispersion Sci. Technol.* **9**, 171–206 (1988).
31. Leikin, S., Parsegian, V. A. & Rau, D. C. A. *Rev. phys. Chem.* **44**, 369–395 (1993).
32. McIntosh, T. J. & Simon, S. A. *Rev. Biophys. biomolec. Struct.* **23**, 27–51 (1994).
33. Ducker, W. A., Senden, T. J. & Pashley, R. M. *Nature* **353**, 239–241 (1991).
34. Derjaguin, B. V. & Landau, L. *Acta Physicochim. URSS* **14**, 633–662 (1941).
35. Verwey, E. J. W. & Overbeek, J. Th. G. in *Theory of the Stability of Lyophobic Colloids* (Elsevier, Amsterdam, 1948).
36. Israelachvili, J. N. & Pashley, R. M. *Nature* **306**, 249–250 (1983).
37. Pashley, R. M. & Israelachvili, J. N. *J. Colloid Interface Sci.* **101**, 511–523 (1984).
38. Henderson, D. J. *Colloid Interface Sci.* **121**, 486–490 (1988).
39. Christenson, H. K. & Yaminsky, V. V. *Langmuir* **9**, 2448–2454 (1993).
40. Derjaguin, B. V. & Churaev, N. V. *Nature* **244**, 430–431 (1973).
41. Franks, F. *Polywater* (MIT Press, Cambridge, MA, 1981).
42. Davenas, E. *et al. Nature* **333**, 816–818 (1988).
43. Maddox, J., Randi, J. & Stewart, W. W. *Nature* **334**, 287–290 (1988).
44. Mysels, K. J. *Langmuir* **8**, 3191–3194 (1992).
45. Hartley, G. S. in *Aqueous Solutions of Paraffin-chain Salts* 60 (Hermann, Paris, 1936).
46. Kjellander, R. & Marčelja, S. *Chem. Phys. Lett.* **120**, 393–396 (1985).
47. Kjellander, R. & Marčelja, S. *Chemica Scripta* **25**, 73–80 (1985).
48. Marinik, S. J., Berkowitz, M. & Berendsen, H. J. C. *Langmuir* **9**, 3122–3131 (1993).
49. Pashley, R. M. *Adv. Colloid Interface Sci.* **16**, 57–62 (1982).
50. McIntosh, T. J., Magid, A. D. & Simon, S. A. *Biochemistry* **26**, 7325–7332 (1987).
51. Lis, L. J., McAlister, M., Fuller, N., Rand, R. P. & Parsegian, V. A. *J. Phys. J.* **37**, 657–666 (1982).
52. Lyle, I. G. & Tiddy, G. J. T. *Chem. Phys. Lett.* **124**, 432–436 (1986).
53. Marsh, D. *Biophys. J.* **55**, 1093–1100 (1989).
54. Marra, J. & Israelachvili, J. N. *Biochemistry* **24**, 4608–4618 (1985).
55. McIntosh, T. J. & Simon, S. A. *Biochemistry* **32**, 8374–8384 (1993).

56. Wiener, M. C. & White, S. H. *Biophys. J.* **61**, 434–447 (1992).
57. McIntosh, T. J., Magid, A. D. & Simon, S. A. *Biochemistry* **28**, 7904–7912 (1989).
58. Persson, P. K. T. & Bergenstahl, B. A. *Biophys. J.* **47**, 743–746 (1985).
59. Iler, R. K. *The Chemistry of Silica* 628, 644 (Wiley, New York, 1979).
60. Iler, R. K. *The Chemistry of Silica* 640 (Wiley, New York, 1979).
61. Frens, G. & Overbeek, J. Th. G. *J. Colloid Interface Sci.* **38**, 376–387 (1972).
62. Vigil, G., Xu, Z., Steinberg, S. & Israelachvili, J. *J. Colloid Interface Sci.* **165**, 367–385 (1994).
63. Scheutjens, J. M. H. M. & Fleer, G. J. *Macromolecules* **18**, 1882–1900 (1985).
64. Iler, R. K. *The Chemistry of Silica* 135 (Wiley, New York, 1979).
65. Meagher, L. J. *Colloid Interface Sci.* **152**, 293–295 (1992).
66. Claesson, P. M., Horn, R. G. & Pashley, R. M. *J. Colloid Interface Sci.* **100**, 250–263 (1984).
67. Shubin, V. E. & Kekicheff, P. *J. Colloid Interface Sci.* **155**, 108–123 (1993).
68. Horn, R. G., Smith, D. T. & Haller, W. *Chem. Phys. Lett.* **162**, 404–408 (1989).
69. Israelachvili, J. N. *J. Colloid and Interface Sci.* **110**, 263–271 (1986).
70. Leermakers, F. A. M. & Scheutjens, J. M. H. M. *J. chem. Phys.* **89**, 3264–3274 (1988).
71. Helfrich, W. *Z. Naturf.* **33a**, 305–315 (1978).
72. Servuss, R. M. & Helfrich, W. *J. Phys. (France)* **50**, 809–827 (1989).
73. Dubois, M. et al. *J. chem. Phys.* **96**, 2278–2286 (1992).
74. Tarazona, P. & Vicente, L. *Molec. Phys.* **56**, 557–572 (1985).
75. Evans, R. & Parry, A. O. *J. Phys. condens. Matter* **2**, SA15–SA32 (1990).
76. Christenson, H. K. *J. phys. Chem.* **90**, 4–6 (1986).
77. Gee, M. L. & Israelachvili, J. N. *J. chem. Soc., Faraday Trans.* **86**, 4049–4058 (1990).
78. Christenson, H. K. *Chem. Phys. Lett.* **118**, 455–458 (1985).
79. Christenson, H. K. in *Modern Approaches to Wettability: Theory and Applications* Ch. 2 (eds Schrader, M. E. & Loeb, G.) 29–51 (Plenum, New York, 1992).
80. Lennard-Jones, J. E. & Dent, B. M. *Trans. Faraday Soc.* **24**, 92–108 (1928).
81. Henderson, D. & Lozada-Cassou, M. *J. Colloid Interface Sci.* **114**, 180–183 (1986).
82. Luzar, A., Bratko, D. & Blum, L. *J. chem. Phys.* **86**, 2955–2959 (1987).
83. Marra, J. *Biophys. J.* **50**, 815–825 (1986); *J. phys. Chem.* **90**, 2145–2150 (1986).
84. Kjellander, R., Marčelja, S., Pashley, R. M. & Quirk, J. P. *J. phys. Chem.* **92**, 6489–6492 (1988); *J. chem. Phys.* **92**, 4399–4407 (1990).
85. Marra, J. *J. Colloid Interface Sci.* **107**, 446–455 (1985); **109**, 11–20 (1986).

ACKNOWLEDGEMENTS. We thank Dr S. Steinberg for her constructive suggestions on an earlier version of the manuscript. J.N.I. was supported by the National Institutes of Health, and H.W. was supported at the Institute of Theoretical Physics at the University of California, Santa Barbara by the NSF and the Göran Gustafsson Foundation.

Structure and mechanism of DNA topoisomerase II

James M. Berger[†], Steven J. Gamblin[†], Stephen C. Harrison* & James C. Wang

Department of Molecular and Cellular Biology and * Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA

The crystal structure of a large fragment of yeast type II DNA topoisomerase reveals a heart-shaped dimeric protein with a large central hole. It provides a molecular model of the enzyme as an ATP-modulated clamp with two sets of jaws at opposite ends, connected by multiple joints. An enzyme with bound DNA can admit a second DNA duplex through one set of jaws, transport it through the cleaved first duplex, and expel it through the other set of jaws.

TYPE II DNA topoisomerases are essential cellular enzymes that function in the segregation of newly replicated chromosome pairs, in chromosome condensation, and in altering DNA superhelicity^{1–4}. All type II enzymes, prokaryotic and eukaryotic, belong to a single family (Fig. 1a)^{5–8}. The eukaryotic enzymes are dimers; the prokaryotic enzymes, bacterial DNA gyrase and DNA topoisomerase IV, are A₂B₂ tetramers. The amino- and carboxy-terminal parts of the eukaryotic enzyme polypeptide are homologous to the gyrase B and A subunits, respectively.

Type II topoisomerases work by cleaving and opening one DNA duplex, passing a second duplex through the opening, and then resealing the break^{9–11}. The DNA cleavage is a transesterification between a pair of tyrosyl residues, one in each half of the dimeric enzyme, and a pair of DNA phosphodiester bonds four base pairs apart³. In this process, the phenolic oxygens of the tyrosines become covalently linked to the phosphoryl groups at the 5' ends of the transiently broken DNA, leaving a pair of hydroxyl groups on the recessed 3' ends. The intervening base pairs then separate, and the pair of tyrosine-linked 5' DNA ends move away from each other, opening a 'gate' in the DNA double helix for the transport of another double-stranded DNA segment. After the second segment has passed through, the enzyme-mediated gate in the first DNA segment closes. A transesterification between the pair of 3' hydroxyl groups and the phosphotyrosyl linkages restores the continuity of the DNA strands and breaks the covalent bonds between the enzyme and the DNA. This mechanism requires that the enzyme undergo dramatic conformational changes, involving displacements as large as 35–40 Å (Fig. 1b).

Type II topoisomerases depend on ATP binding and hydrolysis

for completing the full reaction^{1–3}. Non-hydrolysable analogues of the enzymes, such as the β,γ-imido analogue, AMP-PNP, allow one transport event but prevent enzyme turnover¹². Proteolytic cleavage of the GyrB subunit yields an N-terminal subfragment (the 'ATPase domain'), which binds ATP and AMP-PNP¹³, and a C-terminal subfragment termed B' (refs 14, 15) (Fig. 1a). Binding of AMP-PNP causes the ATPase domain to dimerize^{16,17}. Proteolytic digestion of *Saccharomyces cerevisiae* topoisomerase II with SV8 protease has shown that binding of AMP-PNP leads to significant conformational transitions¹⁸, probably involving (at least in part) dimerization of the ATPase domains. There are three preferred cleavage sites [(A), (B) and (C) in Fig. 1a], but site (A), positioned between the ATPase and B' regions, is sensitive in the absence of AMP-PNP, and site (B), positioned between fragments B' and A', is sensitive in its presence¹⁸. Eukaryotic DNA topoisomerase II bound to a DNA ring can also form a salt-stable complex in the presence of AMP-PNP^{19–22}. The formation of this complex has been attributed to the action of the enzyme as an ATP-modulated clamp, which closes when ATP binds and opens after ATP hydrolysis and release; these ATP-modulated movements are tightly coupled to the transport of one DNA double helix through another^{20–22}.

Type I DNA topoisomerases perform a different reaction, the transient cleavage of a single DNA strand to allow passage of a second strand through the break^{1,2}. The type I enzymes are monomers, and the reaction involves transesterification to a tyrosine, as on each subunit of the dimeric type II enzymes. A 67K (relative molecular mass 67,000) fragment of DNA topoisomerase from *Escherichia coli* has four domains in a ring-like structure²³. A domain interface can separate to give access to the active-site tyrosine. A hinge-and-gate mechanism has been proposed, with the cleaved DNA strand opening up across this interface²³.

[†] Present addresses: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA (J.M.B.); National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK (S.J.G.).

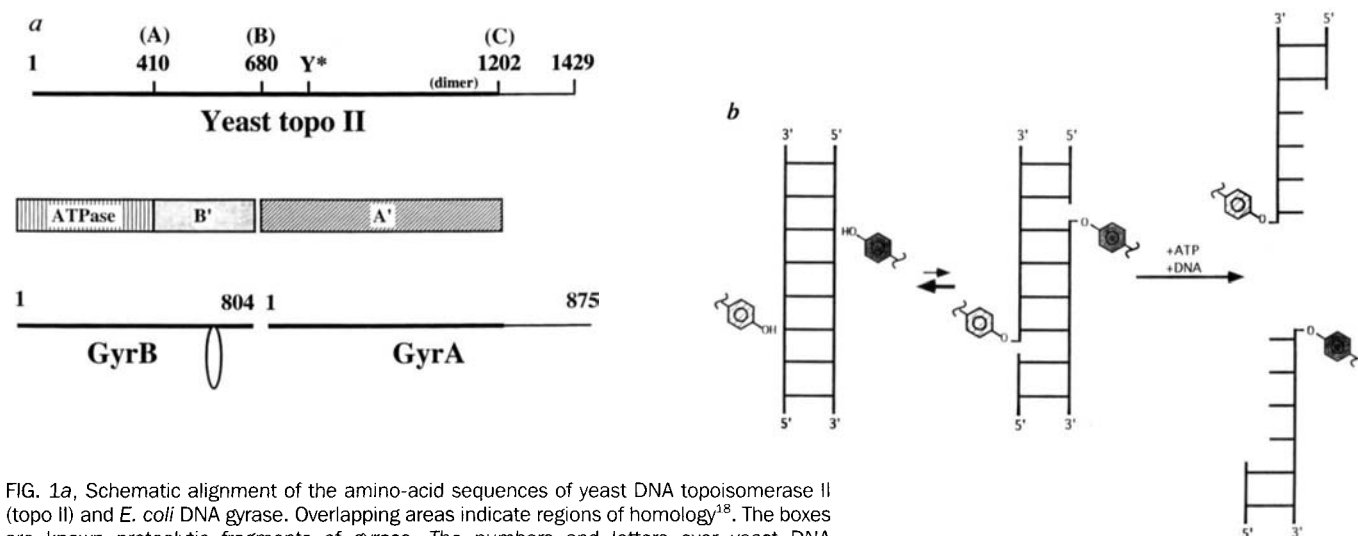


FIG. 1a, Schematic alignment of the amino-acid sequences of yeast DNA topoisomerase II (topo II) and *E. coli* DNA gyrase. Overlapping areas indicate regions of homology¹⁸. The boxes are known proteolytic fragments of gyrase. The numbers and letters over yeast DNA topoisomerase II are the SV8 protease cleavage sites. The dimerization region (defined as the region between residues 1036–1128) and active-site tyrosine positions are labelled dimer and Y*, respectively. b, Diagram of DNA topoisomerase II cleavage reaction. The active-site tyrosines (Tyr 783) in the yeast enzyme are in equilibrium with DNA between cleaved and uncleaved states. ATP binding and perhaps the binding of a second DNA duplex cause the two DNA-linked tyrosines to undergo a significant translocation away from each other to open a gap large enough for the second duplex to pass through.

The ATP-driven transport of one DNA duplex through another has formal analogies with other activities involving displacement of substrate, such as transport through membranes and action of molecular motors. The structure of a 92K fragment of yeast topoisomerase II reported here at a resolution of 2.7 Å (Table 1), shows that the enzyme resembles a clamp with multiple joints and two sets of jaws. The fragment we describe contains residues 410–1,202 of the 1,429-residue polypeptide chain. It lacks the ATPase domain (residues 1–409; Fig. 1a) and a dispensable C-terminal segment (residues 1,203–1,429). The 92K fragment can cleave DNA but cannot execute duplex transport. When combined with the structure of the ATPase domain from *E. coli* GyrB, determined previously¹⁶, our model provides an initial three-dimensional picture of a functional type II topoisomerase.

The monomer

The 92K topoisomerase II fragment (Fig. 2a) has the shape of a flattened crescent with overall dimensions of 120 × 85 × 55 Å. The polypeptide chain folds into two subfragments, denoted B' and A'. B', which corresponds to the B' subfragment of gyrase, contains residues 420–633, whereas A' contains residues 682–1,178 (Fig. 1a). The segment of 48 amino acids between the two subfragments (residues 634–681) is disordered. We have assigned the link between A' and B' (among the possible links to symmetry-related subfragments) based on several criteria. The connection shown produces the shortest distance between the B' carboxy terminus and the A' amino terminus. Moreover, in the dimer (see below), contacts between the A'B' pair connected by the assigned link are significantly complementary, and contain a greater total buried surface area than do contacts in other symmetry-related pairs. Finally, the alternative contacts are between surfaces more accessible to solvent, and are of the kind typically seen as a result of crystal packing. The disordered linker is itself selectively sensitive to protease and corresponds to the discontinuity between distinct subunits in the prokaryotic type II DNA topoisomerases (Fig. 1a).

The B' subfragment consists of two α/β domains (Fig. 2b, top): a five-stranded, mixed β -sheet (residues 420–564 and 606–633), sandwiched by two pairs of α -helices; and a simple β - β - α - β Greek key (residues 565–605) inserted between the fifth strand and final helix of the first domain (Fig. 2c).

The A' subfragment has two snugly docked parts. The N-proximal part of A' (Fig. 2b, middle) contains two distinct domains and a connecting region. The first domain (residues 682–820) has the same fold as the DNA-binding domain of catabolite activator protein (CAP) and histone H5 (Fig. 2d)^{24,25}. The CAP-like domain (Fig. 2b) is augmented in several ways with respect to the domain in CAP itself (Fig. 2d). There is an additional helix at the N terminus, and there are two short helices and a β -hairpin at the C terminus. There are also two long internal loops between the second and third strands and between the third helix and second strand of the domain. Tyrosine 783, the residue that becomes covalently attached to the 5' end of the cleaved DNA chain²⁶, lies

FIG. 2a, Ribbon diagram of yeast DNA topoisomerase II (410–1,202) monomer. The N and C termini of the model are indicated. The two major subfragments are labelled in reference to their homology with prokaryotic gyrase. This figure was generated using MOLSCRIPT¹⁰. b, Exploded view of a. Secondary structure elements are labelled and colour-coded as in a. c, Overlay of secondary structural elements on yeast DNA topoisomerase II (410–1,202) amino-acid sequence. Regions of sequence conservation are shaded grey. Sequence alignment is primarily from (ref. 41), but realignment of the sequence around residues 990–1,076 (based on the three-dimensional structure) revealed a new set of sequence conservations. d, Comparison of the CAP-like domain (residues 702–790, purple) in the A' subfragment with the structure of CAP (green)²⁴ and histone H5 (yellow)²⁵. Secondary structural elements are labelled for the CAP-like domain. The r.m.s. distance between 50 α -carbons of CAP and the CAP-like domain in yeast DNA topoisomerase II is 1.9 Å.

in this region on a loop between the second and third β -strands. The second domain (residues 874–972) is an α - β structure consisting of a four- and a three-stranded mixed β -sheet, backed on one face by three α -helices. The connecting region (residues 821–873 and 973–990) between the two domains contains sequences that link the two domains to each other and to the other part of A'.

The C-proximal part of A' (residues 991–1,178) is largely α -helical (Fig. 2b, bottom). There are three unusually long helices, two of which (A' α 14, A' α 19) make a forked cradle for the first part of A'. One of these two helices (A' α 14) extends to form an antiparallel coiled-coil with the third long helix (A' α 18). The coiled-coil links the forked cradle and the small domain that creates the primary dimer contacts at the 'bottom' of the molecules.

The dimer

The two crescent-shaped monomers form a pair to make a heart-shaped dimer with a large central hole (Fig. 3a). This dimer has overall dimensions of $120 \times 120 \times 55$ Å, and the hole is 55 Å wide at its base, 25 Å wide at the top, and 60 Å long. A large hole has been proposed to exist in the intact yeast topoisomerase II in the presence of AMP-PNP²⁰. The primary dimer contact between subunits occurs at the A'-A' interface. This contact (residues 1,031–1,046 and 1,114–1,130) buries about 1,000 Å² of solvent-accessible surface area in a hydrophobic core flanked by polar interactions. Three cold-sensitive mutations map to this dimer interface²⁷, and appear either to increase the amount of buried hydrophobic surface or to create new salt bridges between monomers, and thus might be expected to stabilize the dimer. Residue

TABLE 1 Crystallographic structure determination

Crystals	Native	CH ₃ HgCl (1)	CH ₃ HgCl (2)	K ₂ PtCl ₄	K ₂ PtCN ₄ (1)	K ₂ PtCN ₄ (2)
Resolution (Å)	15–2.7 Å	14–2.7 Å	14–2.7 Å	14–2.9 Å	8–2.7 Å	14–4.5 Å
No. of crystals	4	1	1	1	1	1
Coverage ($I > \sigma(I)$)						
14–2.7 Å	99.3%	81.0%	77.3%	80.3%	93.4%	59.0%
(at 2.7 Å)	(99.6%)	78.0%	44.5%	79.5%	83.8%	50.2%
R_{merge}^*						
14–2.7 Å	9.0%	8.9%	7.4%	5.9%	7.1%	6.1%
(at 2.7 Å)	56.3%	59.5%	53.0%	33.7% (2.9 Å)	45.6%	13% (4.5 Å)
$R_{\text{iso}}^\dagger$						
14–3.0 Å		16.3%	17.6%	36.5%	15.1%	11.1%
(at 3.0 Å)		21.2%	20.6%	38.7%	19.6%	12.3%
No. of sites		3	3	16	7	3
$R_c^\ddagger$		0.73	0.71	0.93	0.95	0.89
Phasing power§		1.70	1.59	0.74	0.55	0.82
Figure of merit						
14–2.7 Å	Centric 72.7%	Acentric 52.8%	Overall 55.0%			
Refinement						
No. of atoms	No. of reflections (work)	$R_{\text{cryst}}^{\parallel}$	No. of reflections (free)	$R_{\text{free}}^{\parallel}$	r.m.s. bonds	r.m.s. angles
6,997	40,049	22.1%	2,111	27.7%	0.012 Å	2.0°

Data collection. A 92K fragment of *S. cerevisiae* DNA topoisomerase II containing residues 410–1,202 was prepared as follows. *S. cerevisiae* cells bearing a plasmid clone were induced to overexpress topoisomerase II (410–1,202)²⁶. Cells were collected, lysed using glass beads, and purified by several steps of column chromatography in the following order: DEAE (fast-flow, Sigma), S-Sepharose (fast-flow, Sigma), Q-Sepharose (fast-flow, Sigma), Phosphocellulose P-11 (Whatman), and Sephacryl S-300 column (Pharmacia). Purified protein was concentrated to 12 mg ml⁻¹ using a Schleicher and Schuell vacuum dialysis concentrator and then dialysed over 24 h against three changes of buffer (10 mM HEPES, pH 7.4, 270 mM KCl, 0.1 mM 2-mercaptoethanol, 1 μ M leupeptin and 1 μ M pepstatin-A). Crystals of topoisomerase II (410–1,202) were grown at 4 °C from a mixture of 3 μ l protein with 3 μ l reservoir buffer (30 mM sodium cacodylate, pH 6.4, 0.4% PEG 3350, 8 mM MgCl₂ and 290 mM NaCl) using hanging-drop vapour diffusion. The crystals were bipyramids (0.25 $\times$ 0.25 $\times$ 0.15 mm) in space group $P4_32_12$, $a = b = 154.1$ Å, $c = 127.5$ Å. There is one monomer per asymmetric unit (solvent content 70% by volume). The crystals were transferred to a cryo-protectant solution of reservoir buffer plus 5% 2-methyl, 2,4-pentane-diol (MPD) (2–24 h). Crystals were then transferred to a similar solution containing 15% MPD for 2–24 h, and finally to the same solution with 30% MPD for 20 min. At this stage, crystals were flash-cooled to about –160 °C for diffraction measurements. Data were collected at CHESS (beamline F-1, using image plates digitized on a Fuji BAS2000 scanner, and beamline A-1, using a CCD detector), and at SSRL (beamline 7-1, using a 30-cm MAR image plate detector).

Data processing. Data were reduced using the programs DENZO and SCALEPACK³⁴, then processed using the CCP4 suite of programs³⁵. Difference Patterson functions readily identified a single heavy-atom site for the methyl-mercuric chloride derivative (derivative 1). Single isomorphous replacement (SIR) phases generated from this site identified two additional sites in the derivative, whose positions were then confirmed as weaker peaks on the Patterson-Harker sections. Variation of the soak conditions (derivative 2) gave a second mercury derivative, in which the occupancy of one of the sites varied with respect to the other two sites as compared with the first derivative. Phases from these derivatives were used to identify the sites in the platinum derivatives in difference Fourier. As a cross-validation, SIR phases from each of the platinum derivatives were used to identify the heavy-atom sites in the mercury derivatives. Multiple isomorphous replacement (MIR) phases generated from all derivatives yielded an initial electron-density map, which was interpretable only in a few regions of readily discernable secondary structure. **Solvent flattening and refinement.** Solvent flattening and histogram matching were applied to the experimental data using the program dm³⁶. A model was built using this map. As the model-building progressed, atom positions were used to construct improved solvent masks, which were cycled back into density modification to yield improved maps. The process was iterated until it converged with approximately 87% of the residues built. The remaining residues could not be located in these maps. The experimental model was subsequently refined using XPLOR³⁷ from a starting crystallographic R -factor of 49% to its current value of 23.5% over a resolution range from 14–2.7 Å. The free R -factor³⁸ was used as guide for convergence, starting at 52% and ending at 28.9%. The entire refinement process required only two cycles of rebuilding with simulated annealing 'omit' maps³⁹. The first cycle occurred after conventional refinement techniques had already brought the free R -factor to 31.9%.

* $R_{\text{merge}} = \sum_j |I_j - \langle I \rangle| / \sum_j \langle I \rangle$, where I_j is the intensity measurement for the reflection j , and $\langle I \rangle$ is the mean intensity for multiply recorded reflections.

† $R_{\text{iso}} = \sum |F_{\text{ph}} - F_p| / \sum F_p$, where F_{ph} and F_p are the derivative and native structure factors, respectively.

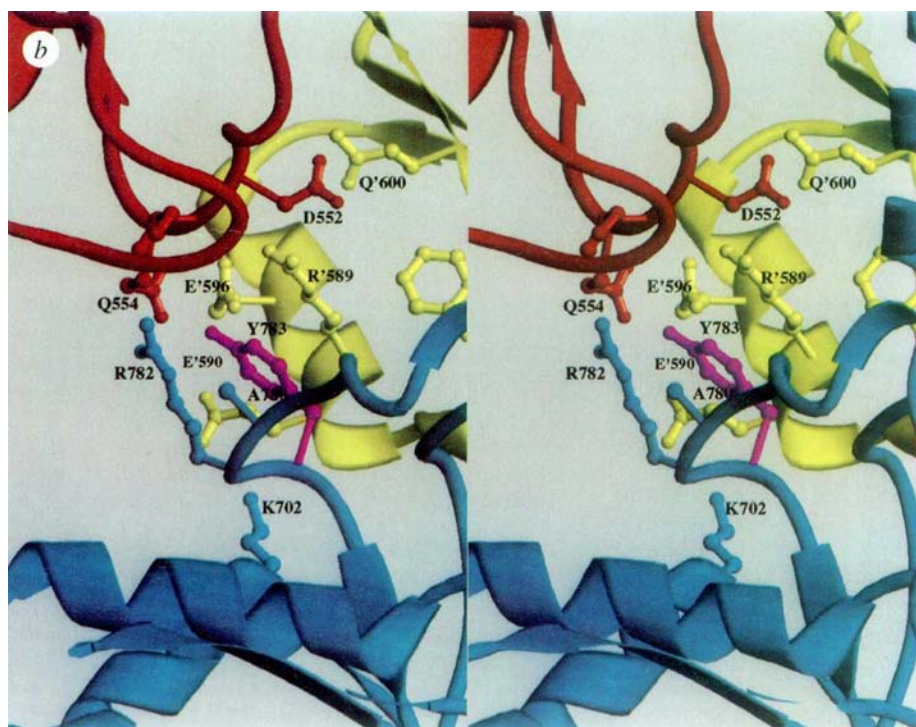
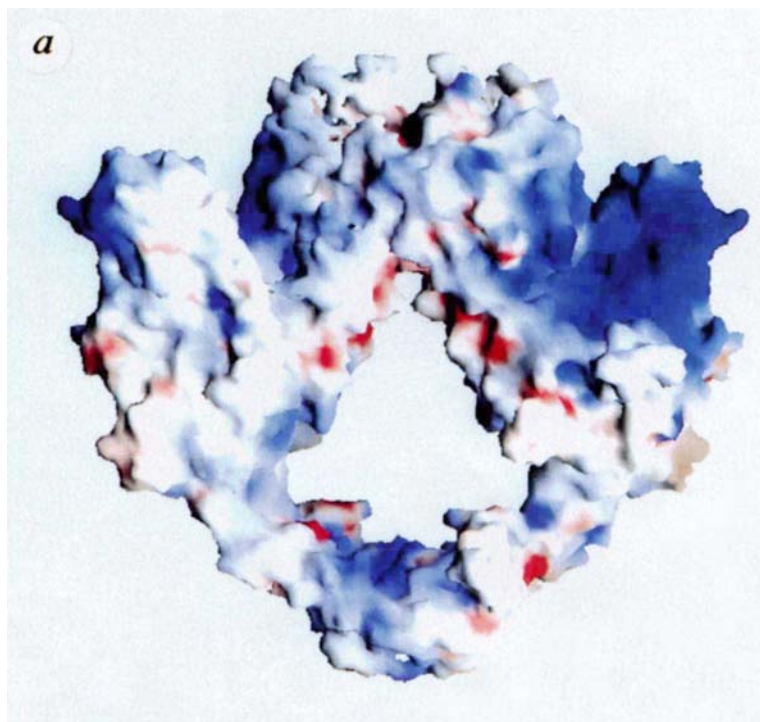
‡ $R_c = \sum |F_{\text{ph}} - F_p| - |F_{\text{hc}}| / \sum |F_{\text{ph}} + F_p|$, where F_{hc} is the calculated heavy-atom structure factor.

§ Phasing power = $\langle F_h \rangle / E$ where $\langle F_h \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error.

|| Non-hydrogen atoms included in the current model.

¶ $R_{\text{cryst, free}} = \sum |F_o| - |F_c| / \sum |F_o|$, where the crystallographic and free R -factors are calculated using the working and free reflection sets respectively ($F_o > 2\sigma(F_o)$). The free reflections (5% of the total) were held aside throughout refinement.

FIG. 3a, Representation of the yeast DNA topoisomerase II (410–1,202) dimer. The B' subfragment is coloured red and the A' subfragment blue. Subfragments belonging to one subunit are shaded more darkly than their counterparts. The active-site tyrosine positions are marked by green spheres and labelled Y*. The primary A' dimer contact is also indicated. This figure was generated using RIBBONS⁴². *b*, Stereo representation of the region around the active-site tyrosines. This view is oriented by rotating the dimer in *a* about the vertical axis by 40° and zooming in on the active site. The B' and A' subfragments that belong to one subunit are coloured red and blue, respectively. The symmetry related B' subfragment is coloured yellow. Residues that belong to different subfragments are colour-coded according to that subfragment; the active-site tyrosine is coloured magenta. Primed and unprimed labels identify residues in symmetry-related subfragments.



1,121, which imparts resistance to the *ccdB* toxin when mutated to alanine from arginine in the homologous *E. coli* GyrA²⁸, also maps to the upper surface of this interface. The V-shaped A'–A' dimer in turn receives a B'–B' dimer at its open end; this protein arch caps the hole between the two A' domains. This arrangement of the pair of monomers is based on the linker assignment discussed earlier, and has implications for the catalytic mechanism.

The B'–B' dimer contact is less extensive than the A'–A' interface, but is still substantial, burying over 800 Å² of solvent-accessible surface area. The two B' subfragments curl about each other, with each of the smaller domains contacting the larger

domains of the other subfragment, and a small channel runs between the two contacts. The B'–B' contact juxtaposes several nearly invariant motifs in the B' sequence across the dyad, placing these regions near the 'top' of the molecule (Fig. 3a). Although some of the contacts involve relatively variable residues, the most intimately associated region is quite well conserved. This region contains a small hydrophobic patch buried in the middle of each leg of the B'₁–B'₂ arch (the subscripts 1 and 2 denote the two individual protomers in a dimer), and includes residues Ile 540, Leu 551, Leu 557 (B'₁) and Tyr 602 (B'₂), which are all conserved except Tyr 602. Replacement of this tyrosine residue by arginine in

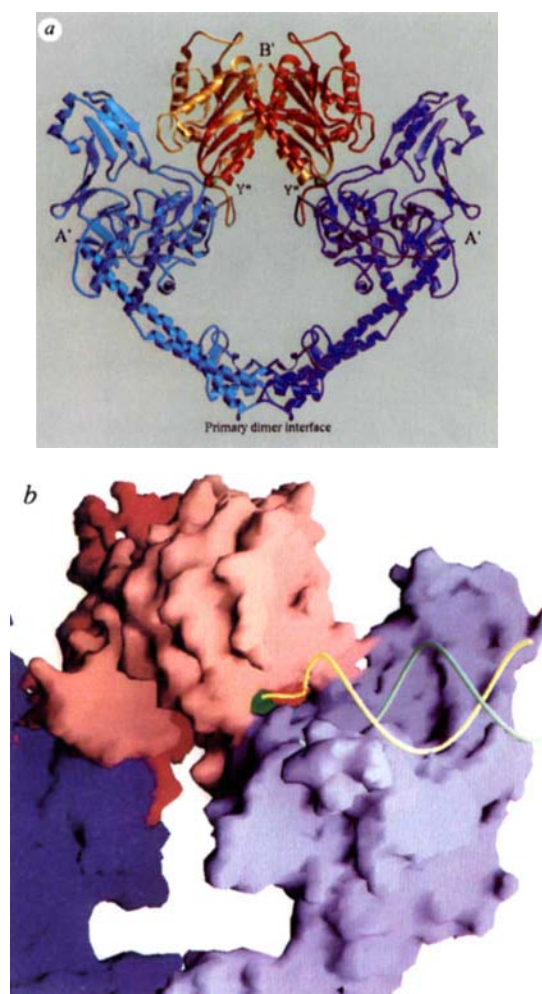


FIG. 4a, Surface representation of the yeast DNA topoisomerase II (410–1202) dimer coloured by electrostatic potential. Regions of basic potential ($>12 k_B T/e$) are in blue, acidic ($<12 k_B T/e$) are in red, where k_B is the Boltzman constant, T the absolute temperature, and e the charge of an electron. An ionic strength of 0 and a dielectric constant of 80 for solvent and 2 for protein were used in the calculation. The darkest blue regions correspond to the A' grooves. This figure was generated using GRASP⁴³. b, Model for DNA binding to topoisomerase II. The orientation is similar to that of Fig. 3b, looking obliquely at the A' groove. The B' subfragment is light red, the A' subfragment light blue. The dimer-related subunit is coloured similarly, only in darker shades. The active site is dark green. A representation of a 12 base-pair duplex DNA backbone with a four nucleotide 5' overhang is shown by yellow and light-green tubes. In this model, the 5' phosphate terminates at Tyr 783.

bacterial type II enzymes is accompanied by replacement of the adjacent Met 582 (B'_2) by glutamate or aspartate. Modelling the charged pair shows that it can form a salt bridge and that the aliphatic portion of the arginine side chain can take the place of the tyrosine-ring contacts with the other conserved hydrophobic residues. This hydrophobic patch is also contacted by the aliphatic portion of the conserved Lys 604 (B'_2), which then extends to form a salt bridge with another conserved B'_1 residue, Asp 531. These conservations within a region of precise complementarity lead us to conclude that the B'_1 – B'_2 interaction in our structure is not merely a crystal-packing contact, and that it represents a functionally significant state.

The B'_1 – B'_2 interface diverges at each base of the arch to accommodate the loop from the CAP-like A' domain which contains the active-site tyrosine, Tyr 783. B' –A' interactions occur primarily between the smaller domain of the B' subfragment and the dimer-related, CAP-like domain of the other monomer,

with fewer contacts occurring between B' and A' domains of the same subunit. Electron-density maps suggest that additional residues (from the linker) may be present in the region between A' and B' domains of the same subunit, but the density is too poorly ordered for a reliable trace.

The dimeric state in the crystals places the active-site tyrosine in an unusual environment. This residue projects from a loop between the second and third strands of the CAP-like domain, and nestles in a cul-de-sac formed by the interaction between both B' domains and the A' domain (Fig. 3b). Thus each active site is effectively shielded from the central hole and is accessed only by a narrow tunnel, which begins at the edge of the B'_2 – B'_1 –A'₂ intersection. The distance between dimer-related, active-site tyrosines is more than 27 Å, a distance too large to represent the pre-cleavage state of the bound DNA duplex.

DNA binding

The domains in the first part of A' create a semicircular groove of 20–25 Å, which funnels down to the tunnel leading to the active site of DNA cleavage. Electrostatic calculations for the entire dimer show that this region has a strong positive potential (Fig. 4a). We have modelled duplex DNA into this groove so that a four-nucleotide 5' overhang extends from the duplex into the tunnel (Fig. 4b). In this arrangement, the homologue of the 'recognition helix' in the CAP-like domain lies in the DNA major groove. Moreover, the loop containing Tyr 783 is the homologue of a backbone-contacting loop (the 'winged' motif) in histone H5²⁵. The recessed 3' end of the duplex lies near the turn of the HTH motif of the CAP-like domain. This HTH region contains five strongly conserved residues (Tyr 735–Glu 739), and its counterpart in *E. coli* gyrase is also a mutational hotspot for resistance to drugs that stabilize the cleaved state of DNA²⁹. The grooves in the first part of A' can accommodate about 10–12 base pairs of duplex. When taken with the middle four bases, this estimate agrees with DNase protection experiments for the size of the footprint of eukaryotic DNA topoisomerase II (25–30 base pairs)^{3,30}. The model for DNA binding is also consistent with rotary shadowing electron microscopy of GyrA–DNA complexes³¹; in these images, GyrA appears as a 'V' shaped molecule, with DNA crossing the open end of the 'V'.

The ATPase domain

The ATP domain (residues 1–409) must connect to the N terminus of the present model, and it is known to dimerize in response to binding of ATP or AMP–PNP^{16,17}. To account for both of these requirements, the ATPase-domain dimer must lie on 'top' of the current model, above the B' subfragments. The GyrB ATPase dimer has a large notch at the dimer interface, closed off into a hole by the conformation of its C-terminal peptides¹⁶. In the intact protein, the C terminus of the ATPase domain is instead covalently joined to B', and we believe that the notch may present surfaces for docking against the B' subfragment, and perhaps for interaction with DNA (Fig. 5).

The existence of different conformational states in response to nucleotide binding has been proposed for yeast DNA topoisomerase II, on the basis of proteolytic mapping and DNA-binding experiments^{18,20–22}. In solution, the purified GyrB subunit is monomeric¹⁷, suggesting that interaction across the B'–B' dimer interface of the crystal structure is weak in the absence of AMP–PNP. In an intact eukaryotic DNA topoisomerase II, B'–B' must also come apart readily to allow DNA to enter. Thus it seems plausible that our structure is related to the ATP-bound state of the enzyme, with the ATPase domains associated across the two-fold axis 'above' their B' partners.

Mechanistic implications

Topoisomerase II performs a cycle of DNA cleavage, transport and religation, coupled to ATP binding and hydrolysis. If DNA binds in the grooves as suggested, with the B' subfragments closed together against it, then our structure represents a state of the

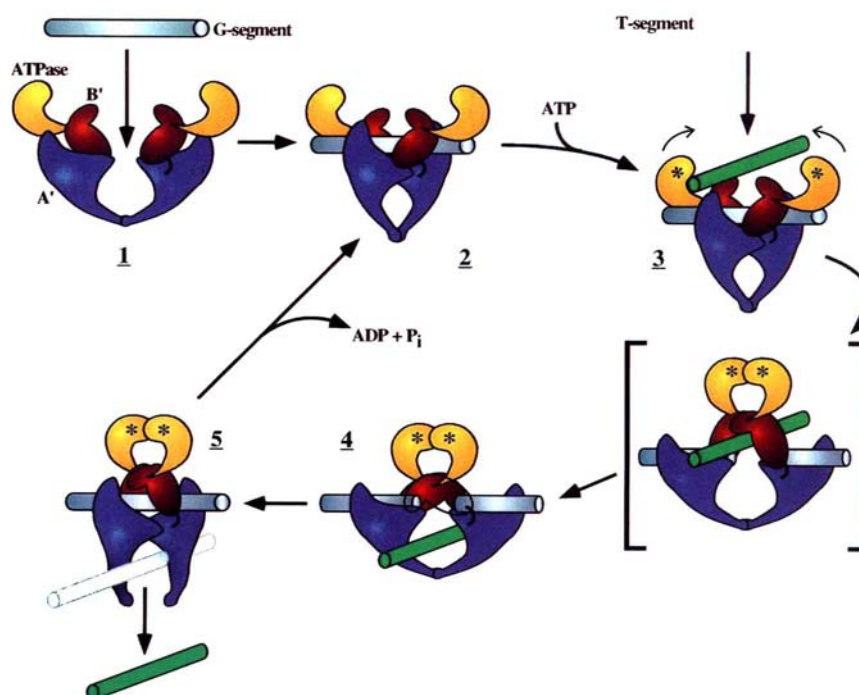


FIG. 5. A molecular model for the catalytic reaction of topoisomerase II. The ATPase domain, B' and A' subfragments are coloured yellow, red and blue, respectively. The G-segment DNA (containing the DNA gate) is grey, and the transported T-segment is green. In (1) an unliganded enzyme first binds the G-segment, inducing a conformational change shown in (2). Upon binding of ATP (represented by asterisks) and a T-segment (3), a series of conformational changes occur in which the G-segment is split by the A' subfragments as they separate from each other. Concomitantly, the ATPase domains dimerize, and the T-segment is transported through the

break and into the central hole (4). The B' subfragment in front is uncoloured in (4) to permit visualization of the DNA behind it. For clarity, the DNA transport step is shown to proceed through a hypothetical intermediate (brackets). Following transport, the G-segment is resealed and the T-segment released from the enzyme through the opening of the dimer interface between A' subfragments (5). The interface between the two A' subfragments again dimerizes, and ATP is hydrolysed and released to regenerate the starting state (2).

protein with ATP bound and the DNA gate open. Although we have neither the ATPase domains nor DNA in our crystals, we can support this conclusion as follows. Our structure shows that yeast DNA topoisomerase II is topologically a ring, with the primary A'-A' dimer contact forming one closed surface and the B'-B' dimer another surface. Previous experiments led us to anticipate this result, and to predict that the molecule would have a hole large enough to 'store' a single DNA duplex²⁰⁻²². To transform the conformation seen in our crystals into one that can bind a DNA duplex across both active site tyrosines, at least two changes must occur. One such change is that the active-site tyrosines, which are approximately 27 Å apart, must move 35-40 Å towards and past each other, to achieve appropriately staggered positions for the 5' ends of the duplex (Fig. 1a). We imagine at least one of several different conformational shifts: a major flexion of the coiled-coil domain, a significant rearrangement of the A'-A' dimer interface, or a rocking of the first A' section within the helical cradle of the second A' section. A second structural change required for DNA binding is expulsion of the B' subfragments from the wide end of the A'-A' 'V'-shape. We propose that ATP hydrolysis and release allow the ATPase domains to separate and also induce the B' subfragments to move away from the dyad. Hinge motions about the two segments of polypeptide chain concerning ATPase to B', and B' to A', could expose the former and protect the latter, as observed in SV8 proteolysis experiments¹⁸.

These structural features, and earlier biochemical observations^{20,22}, lead us to the following model (Fig. 5). In the fully unliganded state (1 in Fig. 5), the enzyme is in an 'open' conformation with the A' subfragments spread apart, as suggested by our structure and by electron microscopy of prokaryotic GyrA³¹. Binding of the DNA segment to be cleaved (the 'gate' or 'G-segment') (2 in Fig. 5) induces a conformational change to

bring the A' subfragments together and the active-site tyrosines into their 'attack' positions. The B' subfragments, perhaps in conjunction with the unliganded ATPase domains, may assist in this change. Indeed, GyrA dimer by itself does not cleave DNA, but GyrA in the presence of the gyrase B' fragment does¹. The yeast DNA topoisomerase II 92K fragment, homologous to B' plus the functional part of GyrA, can also cleave DNA (J.M.B. and J.C.W., unpublished results).

Binding of ATP induces dimerization of the ATPase domains (3 in Fig. 5) and initiates a conformational cascade (4 in Fig. 5). As the ATPase domains dimerize, they capture a weakly held second duplex (the T-segment), pushing it into the interior of the enzyme and through the DNA gate. Steric repulsion by the captured T-segment might aid in separating the cleaved G-segment. Electrostatic calculations also show that, apart from the A' grooves, the most positively charged region is the A'-A' dimer interface. The ATPase domains themselves are relatively acidic in comparison (except for the small interior hole), and therefore the transported T-segment may be drawn through the cleaved G-segment and towards the dimer interface by an electrostatic potential gradient. Furthermore, the B' subfragments may continue the directed transport of the second duplex as they themselves dimerize; at this point the transported duplex is stored in the central hole (4 in Fig. 5).

A DNA duplex threaded into an AMP-PNP-closed topoisomerase (presumably through the large central hole) can still be cleaved²⁰. This result indicates that the B' subfragments in an enzyme with dimerized ATPase domains can separate to allow the threaded DNA duplex to move from the central hole to the A' grooves and active site tyrosines. We know of no published biochemical evidence to suggest that B' subfragments interact with each other, as seen in our structure. As summarized in the

preceding section, however, the intimacy of the contacts in this region and the conservation of many participating residues suggest that the B'-B' interface is not just a crystallization-induced contact. Furthermore, B'-B' dimerization would both block re-entry of the transported duplex back through the split duplex and cover the active site tyrosine-DNA region to prevent hydrolysis of the covalent phosphotyrosine linkage.

The state with the DNA cleaved and separated is probably transient, as the covalent enzyme-DNA intermediate appears only as a minor species when the enzyme-DNA complex is exposed to a protein denaturant³². We therefore expect that the open state will collapse quickly after the transport of the second duplex to a conformation in which the cleaved DNA will religate. How does the transported duplex now exit? Earlier biochemical studies suggested that the A'-A' interface opens to allow its escape²². We suggest that, after the transported DNA enters the central hole, the A' subfragments move towards one another to allow religation of the cleaved DNA. The accompanying constriction of the central hole destabilizes the A'-A' interface sufficiently to open and expel the stored duplex (5 in Fig. 5). Further biochemical evidence in support of this model, using the structural information reported here to create a disulphide-linked A'-A' interface that can be unlinked by a sulphhydryl reducing agent, will be reported elsewhere.

The ring-like topoisomerase II is grossly similar to the prokaryotic topoisomerase I, in that hinge motions on one side of the ring can open up interfaces on the opposite side. The monomeric topoisomerase I has a single active site at such an interface and cleaves a single DNA strand; the dimeric topoisomerase II has two

active sites and cleaves both DNA strands. The only detailed similarity we have noticed is the presence of CAP-like folds in both enzymes. The *E. coli* topoisomerase I contains two domains with CAP-like elements, related by a crude local two-fold axis^{23,33}. One of these elements is in 'domain IV'; the other, more irregular, is in 'domain III' and contains the active-site tyrosine at a position roughly similar to the location of Tyr 783 in the CAP-like domain of yeast topoisomerase II.

In the model shown in Fig. 5, the topoisomerase II dimer has multiple joints, constructed such that the alternating events of breaking and reforming contact surfaces at opposite ends are linked mechanically to several hinge segments. There are hinges between the ATPase domain and the B' subfragments, between the B' and A' subfragments, and within the A' subfragment itself. ATP binding drives formation of the ATPase domain contact surface, and this event in turn facilitates the opening of the DNA gate. Linking association and dissociation of the ATPase domains to ATP binding and hydrolysis appears to have two consequences. It imposes directionality on the transport of DNA from one side of the enzyme to the other, and it couples passage of this DNA duplex through another, held within the enzyme, to an irreversible chemical event. The enzyme also uses an ATP-driven conformational change to overcome the activation barrier for transport. Moreover, the coupling of B-fragment dimerization and G-segment separation ensures that two dimer interfaces hold the complex together whenever a chromosomal break is formed by the topoisomerase. A more detailed analysis of this remarkable mechanochemical device will require structural and biochemical studies of its numerous other states. □

Received 16 October; accepted 17 November 1995.

- Gellert, M. A. *Rev. Biochem.* **50**, 879-910 (1981).
- Wang, J. C. A. *Rev. Biochem.* **54**, 665-697 (1985).
- Hsieh, T. in *DNA Topology and its Biological Implications* (eds Cozzarelli, N. R. & Wang, J. C.) 243-263 (Cold Spring Harbor Laboratory Press, New York, 1990).
- Yanagida, M. & Sternberg, R. in *DNA Topology and its Biological Implications* (eds Cozzarelli, N. R. & Wang, J. C.) 299-320 (Cold Spring Harbor Laboratory Press, New York, 1990).
- Lynn, R., Gaeffer, G., Swanberg, S. L. & Wang, J. C. *Science* **233**, 647-649 (1986).
- Uemura, T. & Yanagida, M. *EMBO J.* **5**, 1003-1010 (1986).
- Wyckoff, E., Natalie, D., Nolan, J. M., Lee, M. & Hsieh, T. J. *molec. Biol.* **205**, 1-14 (1989).
- Huang, W. M. in *Advances in Pharmacology* (ed. Liu, L. F.) 201-225 (Academic, San Diego, 1994).
- Brown, P. O. & Cozzarelli, N. R. *Science* **206**, 1081-1083 (1979).
- Liu, L. F., Liu, C. C. & Alberts, B. M. *Cell* **19**, 697-707 (1980).
- Gellert, M., Fisher, L. M., Ohmori, H., O'Dea, M. H. & Mizuuchi, K. *Cold Spring Harb. Symp. quant. Biol.* **45**, 391-398 (1981).
- Sugino, A., Higgins, N. P., Brown, P. O., Peebles, C. L. & Cozzarelli, N. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4838-4842 (1978).
- Adachi, T., Mizuuchi, K., Menzel, R. & Gellert, M. *Nucleic Acids Res.* **12**, 6389-6395 (1984).
- Brown, P. O., Peebles, C. L. & Cozzarelli, N. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6110-6114 (1979).
- Gellert, M., Fisher, L. M. & O'Dea, M. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6289-6293 (1979).
- Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A. & Dodson, G. *Nature* **351**, 624-629 (1991).
- Ali, J. A., Orphanides, G. & Maxwell, A. *Biochemistry* **34**, 9801-9808 (1995).
- Lindsley, J. E. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10485-10489 (1991).
- Osheroff, N. J. *biol. Chem.* **261**, 9944-9950 (1986).
- Roca, J. R. & Wang, J. C. *Cell* **71**, 833-840 (1992).
- Linsley, J. E. & Wang, J. C. *Nature* **361**, 749-750 (1993).
- Roca, J. & Wang, J. C. *Cell* **77**, 609-616 (1994).
- Lima, C. D., Wang, J. C. & Mondragon, A. *Nature* **367**, 138-145 (1994).
- Schultz, S. C., Shields, G. C. & Steitz, T. A. *Science* **253**, 1001-1007 (1991).
- Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L. & Sweet, R. M. *Nature* **362**, 219-223 (1993).
- Worland, S. T. & Wang, J. C. *J. biol. Chem.* **264**, 4412-4416 (1989).
- Thomas, W., Spell, R. M., Ming, M. E. & Holm, C. *Genetics* **128**, 703-716 (1991).
- Berner, P. & Couturier, M. J. *molec. Biol.* **226**, 735-745 (1992).
- Yoshida, H., Bogaki, M., Nakamura, M. & Nakamura, S. *Antimicrob. Agents Chemother.* **34**, 1271-1272 (1990).
- Lee, M. P., Sauder, M. & Hsieh, T. J. *biol. Chem.* **264**, 21779-21787 (1989).
- Kirchhausen, T., Wang, J. C. & Harrison, S. C. *Cell* **41**, 933-943 (1985).
- Sander, M. & Hsieh, T. J. *biol. Chem.* **258**, 8421-8428 (1983).
- Murzin, A. *Curr. Opin. Struct. Biol.* **4**, 441-449 (1994).
- Otwiński, Z. & Minor, W. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 556-562 (SERC Daresbury Laboratory, Warrington, 1993).
- Collaborative Computational Project, No. 4 *Acta crystallogr.* **D50**, 760-763 (1994).
- Cowtan, K. A. *CCP4 Density Modification Package* (The CCP4 Suite of Programs, 1994).
- Brünger, A. T. *XPLOR (Version 3.1): A System for X-ray Crystallography and NMR* (Yale Univ. Press, 1993).
- Brünger, A. T. *Nature* **355**, 472-475 (1992).
- Brünger, A. T. A. *Rev. phys. Chem.* **42**, 197-223 (1991).
- Kraulis, P. J. *appl. Crystallogr.* **24**, 924-950 (1991).
- Caron, P. & Wang, J. C. *Advances in Pharmacology* (ed. Liu, L. F.) 271-297 (Academic, San Diego, 1994).
- Carson, M. J. *molec. Graphics* **5**, 103-106 (1987).
- Nicholls, A., Sharp, K. A. & Honig, B. H. *Proteins Struct. Funct. Genet.* **11**, 281-296 (1991).

ACKNOWLEDGEMENTS. We thank other members of our research groups for discussions and for help in experimentation, and staff at CHESS and SSRL for help in data collection. This work, a contribution from the Harvard Joint Program in Chemical Biology, was supported by grants from NIH and NSF and by the Howard Hughes Medical Institute. S.C.H. is an investigator in the Howard Hughes Medical Institute.

Discovery of a nearby isolated neutron star

Frederick M. Walter*, Scott J. Wolk* & Ralph Neuhauser†

* Earth and Space Sciences Department, State University of New York, Stony Brook, New York 11794-2100, USA

† Max-Planck-Institut für Extraterrestrische Physik, 85740 Garching, Germany

OUR Galaxy should contain between a hundred million and a billion neutron stars; this follows from an extrapolation of pulsar birthrates¹ and from the number of supernovae required to account for the heavy-element abundances in the Milky Way². Only about 600 pulsars are known³; some fraction of neutron stars will not be beaming in our direction, and therefore will not be detectable as pulsars, and some will be sufficiently old that they have stopped pulsating. How many old neutron stars exist is unknown, but based on the above numbers, about 2,000 isolated ones (not in binary systems) should be detectable as hot thermal sources, emitting X-rays as their surfaces cool or as they accrete gas from the interstellar medium⁴⁻⁷. To date, however, evidence for only one has been presented⁸, and its identification is ambiguous. Here we show that a bright soft-X-ray source has the properties of an isolated neutron star at a distance of about 100 parsecs. Although this confirms our present understanding of how isolated neutron stars should appear, it also highlights the significant problem of accounting for the absence of the others that should be visible.

The X-ray source was identified in data obtained from the satellite Rosat⁹⁻¹¹. The source was detected in the Rosat All Sky Survey (RASS), the 6,332-s Position Sensitive Proportional Counter (PSPC) exposure RP200497, and a 17,800-s High Resolution Imager (HRI) observation RH400612. We analysed the RASS data using EXSAS software, and the pointed (PSPC/HRI) data using the RX package¹². In the PSPC data, taken on 11 and 16 October 1992, the source lies 14.6 arcmin off-axis. After subtracting the interpolated background (determined within an annulus of inner and outer radii of 150 and 450 arcsec, respectively), the net count rate is 3.62 ± 0.03 counts s⁻¹. The net flux is $(1.46 \pm 0.02) \times 10^{-11}$ erg cm⁻² s⁻¹ (0.11–2.4 keV); unidentified X-ray sources of this flux are rare. The HRI image, taken on 7–8 October 1994, is centred on the source position. We extracted the HRI counts from within a 1-arcmin-radius circle; the background level was extrapolated within an annulus of inner and outer radii of 1 and 3 arcmin, respectively. The count rate is 0.56 ± 0.005 counts s⁻¹. The source was also detected in the RASS between 14 September and 2 October 1990, at a count rate of 3.67 ± 0.15 counts s⁻¹. The effective exposure time was 238.7 s in 12 scans.

This source had earlier been detected in the Einstein Observatory slew survey as 1ES1853–379¹³, at 0.86 ± 0.42 counts s⁻¹ (with background box 2: 10 photons detected in an 11.2-s exposure). There was no Einstein pointing at this position.

The position of the HRI source is right ascension 18 h 56 min 35.35 s, $-37^\circ 54' 32''$ ($\pm 1.5''$). The statistical uncertainty is ± 0.1 arcsec, but systematic uncertainties in the boresight correction can reach 10 arcsec in rare cases. The HRI data were processed using SASS version 7.5, which corrects the 0.4° roll angle error present in data processed with earlier versions of SASS (J. Chen, personal communication). We used other X-ray sources in the HRI field to calibrate the boresight offset. Of the six other X-ray sources identified in this region, only one has an obvious optical counterpart on the Digital Sky Survey images. That source is coincident, within 0.7 ± 2.6 arcsec, with the star GSC 7916 0050 ($V = 13$ mag). We conclude

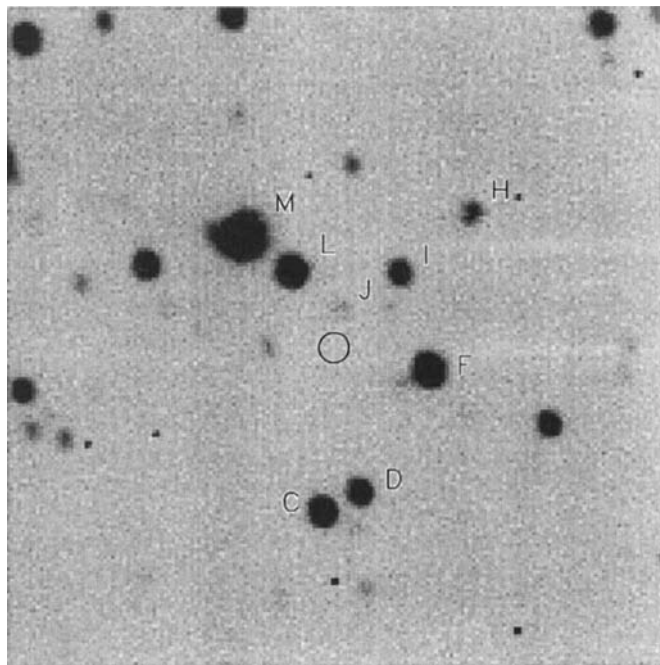


FIG. 1 The central region of the V-band image from the CTIO 0.9-m telescope. The radius of the HRI error circle (at the centre of the image) is 1.5 arcsec. North is up and east is to the left. The grey-scale has been stretched to bring up the fainter stars. The image was reduced using the IRAF routine CCDPROC. Aperture photometry was performed using APPHOT, and the stars were placed on the standard UBVRI system²⁹. The photometric standards agree internally to better than 3%. We reduced the plate solution to the Guide Star Catalog) reference frame by triangulating on HST guide stars in the field. The labelled stars have the following V-band magnitudes: C, 18.8; D, 19.3; F, 18.1; H, 20.4; I, 19.6; J, 22.2; L, 18.4; M, 16.3. Estimated uncertainties are ± 0.03 mag at $V = 19$, and 20% for the faintest targets. None of the stars, to $V = 22.2$, are particularly blue. Only star M was detected in our U-band image, to a limiting magnitude of 19.6 ± 0.3 . The colours of most of the stars are consistent with those of lightly reddened K giants.

that the uncertainty in the boresight correction is unlikely to exceed ± 2.0 arcsec.

There is no obvious optical counterpart at the HRI position in the Palomar Observatory Sky Survey prints or on the Digital Sky Survey image: were the counterpart a hot white dwarf or an active cool star, it would be expected to be brighter than about $V = 14$ mag ($f_X/f_V < 1$), where f_X/f_V is the ratio of the flux in the X-ray band to that in the visible band. To extend the search for the optical counterpart, we obtained images of this field in the Bessell B, V, R and I bands using the CTIO 0.9-m telescope on 28 May 1993, and a single deeper U-band image on 31 May 1994. Seeing was better than 2 arcsec. Exposure times were 10 and 7 min for B- and V-band images, respectively. A pair of 5-min exposures were used for R and I bands. On the second night, only a single 15-min Johnson U-band image was obtained.

The central region of the V-band image is shown in Fig. 1. There is no optical counterpart within the X-ray error circle. The nearest star visible on this image, J, with a V-band magnitude of 22.2, is 4 arcsec north of the X-ray position. From the fluctuations in the background, we estimate the limiting V-band magnitude to be 23.1 (3σ). The f_X/f_V ratio exceeds 7,000.

We fit various spectra to the PSPC data using the software package XSPEC version 8.5. The best fit was a black body, with a reduced χ^2 of 2.3 for 30 degrees of freedom (Fig. 2). A spectra fit to the RASS data gave identical results, within the errors. The large χ^2_ν may reflect the poorly understood and time-variable PSPC spectra response at low energies¹⁴, or may indicate that the

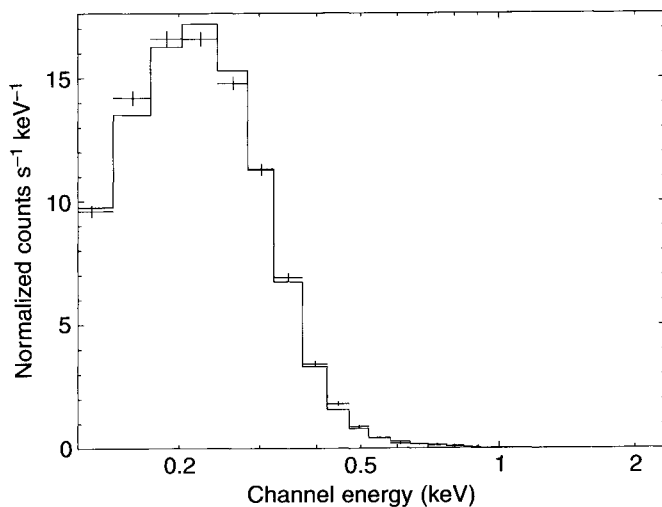


FIG. 2 The best-fit black-body spectrum. We fit pulse height channels 3–32 (0.11–2.23 keV) using XSPEC version 8.5. The best fit was $kT = 57 \pm 1$ eV and $n_H = (1.4 \pm 0.1) \times 10^{20} \text{ cm}^{-2}$. The relatively poor χ^2_ν suggests the formal uncertainties may be too small. Other spectral models yielded significantly larger values of χ^2_ν .

intrinsic spectrum is not a black body. Other spectral models gave larger values of χ^2_ν .

We searched for evidence of emission from this source at other wavelengths. The source does not appear in the HEAO-1 A1 LED catalogue¹⁵. It is not a detectable extreme ultraviolet (EUV) source, appearing in neither the Rosat WFC 2RE catalogue¹⁶ nor the EUVE bright source list¹⁷. The EUVE scanners observed this region for 1,001 s; the EUVE Lexan filter count rate is $< 0.066 \text{ counts s}^{-1}$ (3σ). The EUV limits can be attributed to the interstellar absorption. The position is contained within two GRO 2B catalogue γ -ray-burst positions. The X-ray source lies 4.6° from the γ -ray object 2B 920222 (6.3° error radius), and 1.4° from 2B 921128 (4.4° error radius). There is no known radio pulsar near these coordinates from the Parkes pulsar survey, with a flux limit of 4 mJy at 430 MHz (S. Johnston, personal communication). In a 330-s VLA snapshot of this region, Perlman *et al.*¹⁸ attained a 5σ upper limit of 0.6 mJy at 5 GHz. We also examined the IRAS ADDSCAN data, which are high signal-to-noise and one-dimensional, as well as the two-dimensional FRESCO and HIRES frames of the region. We found no source within one arcminute of the X-ray position, and we were able to put limits on the flux from this region as 0.12, 0.11, 0.20 and 0.36 Jy at wavelengths of 12, 25, 60 and 100 μm respectively.

The X-ray source position is projected on the east tail of the R. Corone Australis (R CrA) molecular cloud. Wang¹⁹ estimates the total gas column density in the cloud at this position to be $\sim 5 \times 10^{20} \text{ cm}^{-2}$. As this exceeds the extinction measured to the X-ray source, it is unlikely that the X-ray source is seen through the bulk of the molecular gas. Were the source embedded in the molecular cloud, we would expect heating of the cloud to produce an infrared source²⁰; lack of an IRAS detection suggests that the source is foreground to the R CrA cloud. The assumed distance of the R CrA cloud is 120 pc; the X-ray source is probably closer.

The black-body spectrum, extrapolated into the optical, yields unreddened U-, B- and V-band magnitudes of 25.0, 26.4 and 26.7, respectively. The small X-ray column corresponds to the total extinction, $A_V \sim 0.1$ mag. The intrinsic f_X/f_V for the unreddened object is 4×10^5 .

The variability of X-ray sources can give clues to their nature. There is no evidence for X-ray variability in this source. The fluxes derived from the four independent X-ray observations all agree to within the observational uncertainties. Although the PSPC data are variable on all timescales, the cause is probably instrumental.

A fast Fourier transform analysis shows some power at ~ 80 –100 s, as has been seen in other PSPC sources (H. Ögelman, personal communication), and is probably due to the spacecraft wobble. The HRI data are not subject to the instrumental variability induced by the shadowing of the PSPC window support wires. A one-sided Kolmogorov–Smirnov test applied to the arrival times shows that one cannot reject the hypothesis of a constant count rate at the 1σ level. χ^2 -square tests to binned data are also consistent with the absence of variability on timescales from a few seconds to a day. We conclude that there is no evidence for variability at levels greater than given by photon statistics on short timescales ($\sim 1\%$ in an hour), or by our ability to inter-compare different instruments ($< 30\%$) on timescales up to 15 years.

The soft black-body spectrum is suggestive of the hot, optically thick surface of a compact object. The 60-eV temperature is consistent with hot white dwarfs, accreting white dwarfs in cataclysmic binaries, or ultrasoft X-ray sources in X-ray binaries²¹. However, the high value of f_X/f_V greatly limits the possibilities. The low interstellar column, and the fact that the measured column is less than the estimated column depth through the R CrA molecular cloud, rules out any extragalactic source. The extreme softness of the source argues against a disk-accretion source, and the large f_X/f_V ratio excludes most accretion disks. The lack of X-ray variability also excludes many accretion-powered systems. The lack of X-ray or radio pulsations excludes a pulsar (unless the beam is not directed at us). By a process of elimination, we are left with the possibility that this source is a nearby isolated old neutron star.

On the assumption of black-body emission, the observed flux and temperature require an emission area of $\sim 480(D/(100 \text{ pc}))^2 \text{ km}^2$, where D is the actual distance to the source. The upper limit to D of 120 pc suggests that either the source is the hot surface of a cooling neutron star, or that accretion heating occurs over much of the surface of a neutron star, and not only at the magnetic poles. The data do not permit us to discriminate between these two emission models. However, the cooling time of neutron stars is short: neutron stars are expected to cool to 57 eV in $\sim 10^6$ yr, or within ~ 100 yr if direct Urca cooling occurs^{22,23}. It is highly unlikely that this is a 100-year old neutron star at a distance of 100 pc, as a recent supernova at this distance would probably have been noticed. A 10^6 -year-old high-velocity object could have moved a good fraction of a kiloparsec from its birthplace. It has also been suggested that old neutron stars reheat their surfaces²⁴ to about this temperature. Because of its proximity to the R CrA molecular cloud, it is possible that this is an old neutron star accreting from the interstellar medium.

Isolated old neutron stars may accrete out of the interstellar medium, and if they do the bulk of their radiation will be emitted in the EUV and soft X-rays^{7,25,26}. Blaes and Madau⁷ modelled this emission, assuming a black-body shape for the spectrum, finding temperatures $T_{\text{BB}} = 20(M_{10}/f)^{1/4} \text{ eV}$, where $\dot{M}_{10}$ is the mass accretion rate in units of 10^{10} g s^{-1} , and f is the fraction of the neutron star surface upon which accretion occurs. The bolometric luminosities (mostly in the EUV and soft X-rays) are $L_{\text{bol}} = 2 \times 10^{30} M_{1.4} R_6^{-1} \dot{M}_{10} \text{ erg s}^{-1}$, where $M = 1.4 M_{1.4} M_\odot$ and $R = 10^6 R_6 \text{ cm}$.

For $f = 1$ (isotropic accretion), an ambient interstellar density $n_H = 1 \text{ cm}^{-3}$, and a neutron-star velocity v of 5 km s^{-1} with respect to the gas, the Blaes and Madau⁷ model predicts an X-ray luminosity, L_X (0.2–2.4 keV) = $7 \times 10^{31} \text{ erg s}^{-1}$. Given the considerable latitude offered by the unknown parameters f , v , n_H , and D , these parameters can be easily adjusted to match the observed luminosity of $5 \times 10^{31} (D/100 \text{ pc})^2 \text{ erg s}^{-1}$. The data are consistent with the picture of accretion of the ambient interstellar medium by a slowly-moving neutron star at a distance of the order of 100 pc.

Confirmation that this object is a neutron star will require measurement of its parallax. At 100 pc, the Hubble Space Telescope could obtain a parallax and proper motion, permitting a discrimination between the cooling and the accretion heating

models for the X-ray emission. As $L_X \approx n_H$ for accretional heating, long-term X-ray monitoring may provide a probe of the inhomogeneity of the interstellar medium on small scales. X-ray observations with better spectral resolution, together with ultraviolet and optical fluxes, can reveal the spectral energy distribution, and test various predictions. Zampieri *et al.*²⁷ have argued that the spectra of accreting, non-magnetized neutron stars are significantly harder than black bodies. Nelson *et al.*³⁰ have argued that the spectra of accreting, strongly magnetized neutron stars can contain a broadened hard cyclotron emission feature.

We note that of the thousands of isolated old neutron stars expected to be observable with current instrumentation, and despite careful searches²⁸, only two candidates have so far been put forth. Are these isolated old neutron stars far less numerous than had been thought, is their velocity dispersion greater than that of the pulsars, or are there selection effects that conspire against their detection? By studying the current candidates we may be able to devise more efficient search techniques, and determine where this (presumably) dominant class of neutron stars is hiding. Elucidation of the properties of isolated old neutron stars will have important ramifications for many areas of galactic astronomy. For example, observations of such stars could be used to further study the velocity dispersion of neutron stars, with implications for the symmetries of supernova explosions. Detailed study of individual isolated old neutron stars could determine the heating mechanisms (or cooling times), and would help constrain the equation of state of neutron stars. The properties of isolated old neutron stars are very poorly understood theoretically, and the detection of even a single neutron star of this type close enough to study would lead to a great increase in our understanding of the physics of neutron stars. $\square$

Received 19 July; accepted 15 November 1995.

1. Narayan, R. & Ostriker, J. P. *Astrophys. J.* **352**, 222–246 (1990).
2. Amett, W. D., Schramm, D. N. & Truran, J. W. *Astrophys. J.* **339**, L25–L27 (1989).
3. Phinney, E. S. & Kulkarni, S. R. *A. Rev. Astr. Astrophys.* **32**, 591–639 (1994).
4. Ostriker, J. P., Rees, M. J. & Silk, J. *Astrophys. Lett.* **6**, 179–184 (1970).
5. Helfand, D. J., Chanan, G. A. & Novick, R. *Nature* **283**, 337–343 (1980).
6. Treves, A. & Colpi, M. *Astr. Astrophys.* **241**, 107–111 (1991).
7. Blaes, O. & Madau, P. *Astrophys. J.* **403**, 690–705 (1993).
8. Stocke, J. T., Wang, Q. D., Perlman, E. S., Donahue, M. E. & Schachter, J. *Astr. J.* **109**, 1199–1203 (1995).
9. Trümper, J. *Adv. Space Res.* **2**, 241–249 (1983).
10. Pfeffermann, E. *et al.* in *Soft X-Ray Optics and Technology* (eds Koch, E.-E. & Schmahl, G.) 519–557 (Proc. SPIE Vol. **733**, SPIE: Bellingham WA, 1987).
11. Aschenbach, B. *Bull. appl. Opt.* **27**, 1404–1413 (1988).
12. Walter, F. M. in *The Soft X-Ray Cosmos* (eds Schlegel, E. M. & Petre, R.) 263–265 (AIP Conf. Proc. **313**, Am. Inst. Phys., New York, 1994).
13. Plummer, D., Schacter, J., Garcia, M. & Elvis, M. *The EINSTEIN Observatory IPC Slew Survey CD-ROM* (Center for Astrophysics, Cambridge MA, 1991).
14. Snowden, S. L. *et al.* *Status of the PSPC Spatial/Temporal Gain Calibration* (OGIP Calibration Memo CAL/ROS/95-003, NASA Goddard Space Flight Center, Washington, DC, 1995).
15. Nugent, J. *et al.* *Astrophys. J. Suppl. Ser.* **51**, 1–28 (1983).
16. Pye, J. *et al.* *Mon. Not. R. astr. Soc.* **274**, 1165–1193 (1995).
17. Malina, R. F. *et al.* *Astr. J.* **107**, 751–764 (1994).
18. Perlman, E. S. *et al.* *Astrophys. J. Suppl. Ser.* (submitted).
19. Wang, Q. D. in *The Soft X-Ray Cosmos* (eds Schlegel, E. M. & Petre, R.) 301–305 (AIP Conf. Proc. **313**, Am. Inst. Phys., New York, 1994).
20. Lepp, S. & McCray, R. *Astrophys. J.* **269**, 560–567 (1983).
21. White, N. E. & Marshall, F. E. *Astrophys. J.* **281**, 354–359 (1984).
22. Page, D. & Applegate, J. H. *Astrophys. J.* **394**, L17–L20 (1993).
23. Lattimer, J. M., van Riper, K. A., Prakash, M. & Prakash, M. *Astrophys. J.* **425**, 802–813 (1994).
24. Reisenegger, A. *Astrophys. J.* **442**, 749–757 (1995).
25. Colpi, M., Campana, S. & Treves, A. *Astr. Astrophys.* **278**, 161–166 (1993).
26. Blaes, O., Warren, O. & Madau, P. *Astrophys. J.* (in the press).
27. Zampieri, L., Turolla, R., Zane, S. & Treves, A. *Astrophys. J.* **439**, 849–853 (1995).
28. Shemi, A. *Mon. Not. R. astr. Soc.* (in the press).
29. Landolt, A. U. *Astr. J.* **104**, 340–371 (1992).
30. Nelson, R. W., Wang, J. C. L., Salpeter, E. E. & Wasseman, I. *Astrophys. J.* **438**, L99–102 (1995).

ACKNOWLEDGEMENTS. We thank J. Lattimer, R. W. Nelson, H. Ögelman, J. T. Stocke and J. C. L. Wang for useful conversations; E. Perlman for providing the radio flux of 1ES1853-379 before publication; and J. Lim and S. Johnston for forwarding the results of the Parkes pulsar search. The EUVE limiting flux was obtained through the Center for Extreme Ultraviolet Astrophysics' Skyflux request facility. We acknowledge the contribution of Rosat personnel in Germany and at the NASA Goddard Space Flight Center to the success of Rosat. The Rosat project has been supported by the German government and the Max-Planck-Society. This work was supported at SUNY Stony Brook by NASA.

Downflows under sunspots detected by helioseismic tomography

T. L. Duvall* Jr, S. D'Silva†, S. M. Jefferies‡, J. W. Harvey† & J. Schou§

* Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

† National Solar Observatory, National Optical Astronomy Observatories, PO Box 26732, Tucson, Arizona 85726, USA

‡ Bartol Research Institute, University of Delaware, Newark, Delaware 19716 USA

§ Stanford University, HEPL, Stanford, California 94305-4085, USA

SUNSPOTS are areas of cooler gas and stronger magnetic fields in the Sun's photosphere (its 'surface'), but just how they form and are maintained has long been a puzzle. It has been proposed¹ that small vertical magnetic flux tubes, generated deep within the Sun, develop downflows around them when they emerge at the surface. The downflows bring together a large number of flux tubes in a cluster to form a sunspot, which behaves as a single flux bundle as long as the downflows bind the flux tubes together. Until now, however, it has not been possible to test this model with sub-surface observations. Here we use the recently developed technique of travel-time helioseismology² to detect the presence of strong downflows beneath both sunspots and the bright features known as plages. The flows have a velocity of $\sim 2 \text{ km s}^{-1}$, and they persist to a depth of about 2,000 km. The data suggest, however, that the vertical magnetic field can be a coherent flux bundle only to a depth of $\sim 600 \text{ km}$; below this depth it is possible that the downflows hold together a loose collection of flux tubes to maintain the sunspots that we see.

The recent discovery² of how to measure the time taken by acoustic waves to travel inside the Sun is a powerful tool for studying the solar interior. This approach directly measures the travel time of acoustic waves between any point on the solar surface and a surrounding annulus. As a packet of acoustic waves propagates into the solar interior, it refracts owing to the increase in sound speed until it eventually turns around and returns to the surface where it is reflected back into the interior. If the wave-packet encounters a sub-surface structure, such as a flow field, magnetic field, or temperature perturbation (which introduces an inhomogeneity in the wave speed along the path), its travel time will be altered. We introduce a technique to measure the alterations in the travel time at each point on the solar surface, and obtain two-dimensional maps from which we locate and measure the sub-surface structures. We call this technique helioseismic tomography, where, as in seismic tomography³, we are attempting to reconstruct the interior of the object from one-dimensional line integrals.

We have generated travel-time maps using 1,017 full-disk images, sampled once per minute in the Ca⁺ K-line, obtained at the geographical South Pole on 4–5 January 1991. This series was divided into two 8.5-h segments, to enable repeatability to be tested. Each segment was processed using the following procedure: (1) The images were corrected for offset and gain, and the high spatial-frequency information restored to each image⁴. (2) The images were interpolated onto a grid of longitude ϕ and sin λ , where λ is latitude (see Fig. 1). (3) The images were filtered to isolate oscillation signals, and spatially and temporally tapered to zero at the edges. (4) For each time series, motions due to a mean solar rotation were removed at each latitude separately using a Fourier shift. (5) The mean temporal cross-covariance function between each point of the images and all points of the image within a narrow annulus was computed. (6) Signal-to-noise ratio was enhanced at the expense of spatial resolution by averaging the cross-covariance functions in spatial blocks of 7×7 map points. (7) The cross-covariance functions were averaged, after subtracting

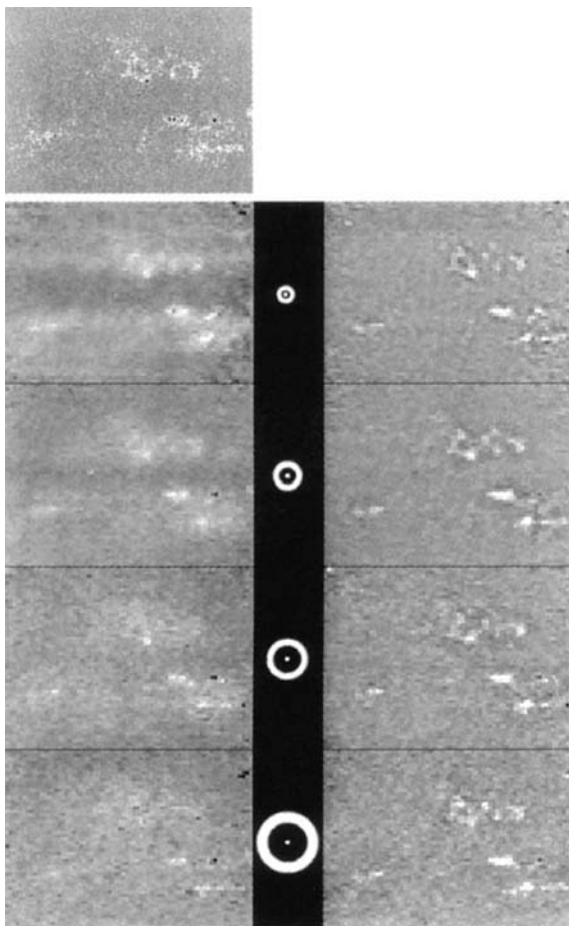


FIG. 1 Upper left, an image of the Sun at 393.3 nm ($\text{Ca}^+ \text{K-line}$) obtained at the South Pole on 4 January 1991. The image has been interpolated onto a grid of longitude (ϕ) and sine of latitude (λ), and only the part corresponding to the subsequent maps is shown. North is to the top and west is to the right. The longitude range covered is $\pm 62.5^\circ$ about the central meridian, and the range in $\sin \lambda$ is ± 0.8 . The small dark spots are sunspots, which are dark because they are cooler than the surrounding surface and are the sites of the strongest magnetic fields measured on the surface. The brightenings seen in the image are called plages. They are also sites of strong magnetic fields, but not as strong as the spots, and are thought to be bright because of heating of the upper atmosphere in these regions. One of the main goals of helioseismic tomography is to understand the sub-surface magnetic, thermal, and flow structure of these surface features. Below this image of the solar surface are maps of the mean and difference travel times. The left-hand column is the mean travel time, and the right-hand column is the difference (outward wave time – inward wave time). Longitude is along the abscissa of each map and $\sin \lambda$ is on the ordinate with the scales being the same as for the solar image. The four maps in each column correspond to the four annuli (see text), with the smallest annulus at the top. The sizes of the annuli are shown between the two columns. The mean maps are displayed as the deviation of the travel time from the mean value for that distance. Grey is zero excess time; white features are shorter-than-average times. The strongest features correspond to a shorter travel time of ~ 0.4 min. For the difference maps, white features correspond to outward-going waves taking a shorter time than inward-going waves. The largest differences correspond to ~ 1 min. Noise has been estimated from the apparently random scatter in a quiet area away from the solar activity. The standard deviation is 0.03–0.04 min for the mean maps, and 0.04–0.09 min for the difference maps, with the larger values for the larger annuli.

tion of a mean time–distance curve, in four ranges of angular distance, or annuli, namely 2.5° – 4.25° , 4.5° – 7° , 7.25° – 10° and 10.25° – 15° . (8) The travel times were measured from the zero crossing of the instantaneous phase of the cross-covariance function determined from a Hilbert transform technique⁵. (9) A ‘mean’ travel-time map was made by averaging the two sets of time measurements, corresponding to the travel time of the wave from the central point to the annulus and *vice versa*, and subtracting this from a map of the average of all travel times over the surface. A ‘difference’ travel-time map was made by differencing the two sets of measurements. The maps for the two 8.5-h segments showed obvious features that rotated with the Sun, and so the two segments were combined after a correction for solar rotation; the resultant maps are shown in Fig. 1.

The mean map indicates how sub-surface structure has altered the local wave speed from the average value. An increased wave speed decreases the travel time, and shows up as a bright region on the mean map. The mean maps in Fig. 1 show that bright regions are correlated with the active regions in that figure. The peak signal in these regions corresponds to sunspots and is roughly -0.4 min for the smallest annulus, and fades away with annulus size. This behaviour could be caused by magnetic fields, which are known to alter the temperature and density of the region they occupy. Lacking a unique solution to describe the internal structure of the magnetic field, we study the effect by introducing local perturbations of temperature and magnetic field separately into a solar model⁶. We compute the ray paths and calculate the change in travel time with and without the perturbation. The travel time decreases by ~ 0.4 min, if the sound speed is increased by 6% in a cylindrical region 5,000 km deep and 0.5° in diameter about the central point. A magnetic field changes the wave speed depending on the orientation of the ray path with the field. Our calculations show that although a horizontal magnetic field can increase the

effective wave speed, it cannot reproduce the observed falloff of the mean signal with increasing annulus size. A vertical magnetic field, on the other hand, can reproduce the observed trend. A signal similar to that observed can be reproduced with a field strength of 2 kG at the surface, increasing to ~ 4 kG at a depth of 600 km. To match the observations, a weaker field would have to extend deeper (Fig. 2c). As the typical field strength of a sunspot is 2 kG, we suggest that the mean signal is largely due to a vertical magnetic field which extends ~ 600 km below the solar surface.

The mean maps also show a latitudinal variation of travel time which fades with annulus size and has no obvious counterpart in the solar activity image. The almost symmetric variation about the solar equator (see Fig. 3) is reminiscent of the latitudinal variation of the wave speed perturbation derived from *p*-mode splittings⁷ and of the latitudinal brightness function at this time of the solar cycle⁸. The mean signal in the bands is roughly -0.1 min for the smallest annulus, and the bands appear to shift in latitude depending on annulus size. A 8% decrease in density scale height (which is the inverse of the density gradient with depth) caused by a vertical magnetic field would explain the magnitude of this signal. A temperature perturbation alone would not explain the band, as it would require a 2% increase in sound speed over the entire latitude band to a depth of 5,000 km. Such an increase in sound speed would result in an ~ 200 K increase in surface temperature, which could have easily been detected. We speculate that these bands and the *p*-mode splitting perturbations share the same origin.

The difference maps detect the presence of any sub-surface structure that alters the wave speed differently when the direction of the ray propagation is reversed. A flow field increases the wave speed if the ray is travelling along the flow, or decreases the wave speed if it is against the flow. Hence the rays travelling from the centre to the annulus would have a different travel time depending

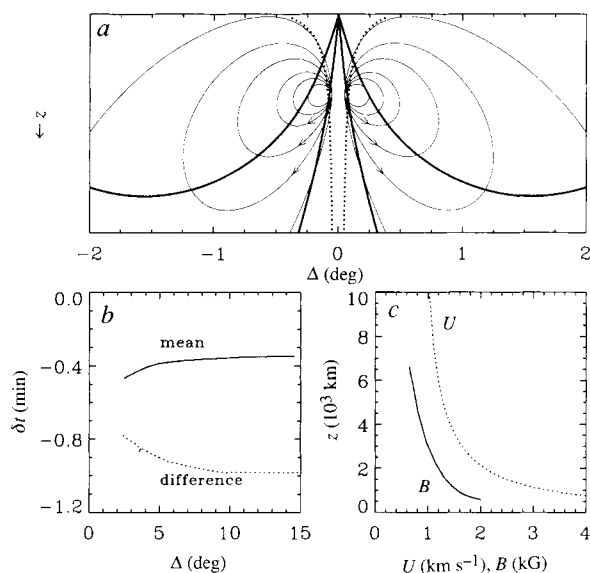


FIG. 2 a, Schematic diagram of Parker's hypothetical downflow⁹ (thin solid lines) around a vertical magnetic flux tube (dots) whose field strength increases with depth (z) from the surface. The horizontal axis is the angular separation Δ , with $1^\circ \approx 12,000$ km at the surface. Thick solid lines are the ray paths which reach the smallest and the largest annulus; the more-vertical rays reach the largest annulus. b, A vertical magnetic field, or a downflow, is introduced in a solar model⁶ within a cylindrical region surrounding the central point, of diameter 0.5° , and variable depth. The mean and difference signals are computed separately for the rays going from the central point to the annulus and back. The solid line is the mean signal as a function of annulus size for a vertical magnetic field 1,000 km deep, whose surface field strength is 2 kG, and increases to 5 kG. The magnetic field region is cooled by 30% as in sunspots, and a region between 1,000 and 2,500 km is heated by 15%, as heat is expected to accumulate beneath the sunspot. The dotted line is the difference signal for a 1 km s^{-1} downflow extending to a depth of 10,000 km. c, The depth z to which a vertical magnetic field should exist plotted against the field strength B for a family of solutions giving the same predicted signal (solid line). The depth to which a downflow should exist to give a difference signal of 1 min in the largest annulus is plotted against the downflow speed U (dotted line). A 2 kG vertical magnetic field 600 km deep, and a 2 km s^{-1} downflow 2,000 km deep, can explain the observed mean and difference signals.

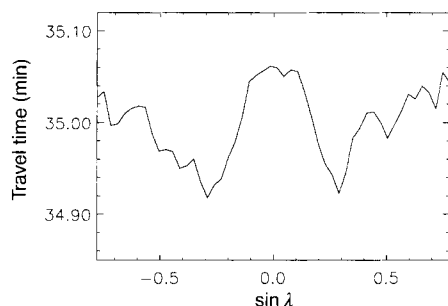


FIG. 3 Travel time (from the mean travel-time map with smallest annulus), averaged over longitude, plotted against sine of latitude λ .

on the direction of the flow. The difference signal therefore specifies the direction of the flow, measures its magnitude, and determines the size of the region to which it is confined. If the outgoing ray from a point has a shorter travel time, then the difference signal is negative and shows up as a bright region on the difference map. Bright regions in the difference maps (Fig. 1) correlate with active regions in Fig. 1, indicating the presence of a flow that is directed outwards from the location of the active regions. The peak difference signal is roughly -1 min in the active regions for the largest annulus, and decreases as the annulus size decreases. Horizontal outward flows are extremely unlikely as calculations demand an extraordinarily large velocity of several kilometres per second spread over several tens of thousand kilometres below the surface. As the rays are largely vertical near the surface (Fig. 2a), a downflow would increase the wave speed of the outgoing ray and decrease that of the incoming ray. The dashed line in Fig. 2b shows the effect of a 1 km s^{-1} vertical downflow confined to a cylindrical region of 0.5° in diameter around the central point with a depth of 10,000 km. A larger flow of 2 km s^{-1} would need to persist only to $\sim 2,000$ km below the surface (Fig. 2c). At the central point, the rays that reach the larger annulus are closer to vertical and stay within the flow longer than those that reach the smaller annulus. This explains why the maximum difference signal is observed for the largest annulus. At the surface, measured horizontal flows surrounding old sunspots are opposite to the inflows expected from sub-surface downflows. This suggests that the surface flows are shallow counterflow phenomena.

We have measured the perturbations in acoustic-wave travel time caused by sub-surface structures, and have qualitatively explained the measurements in terms of downflows and magnetic

fields beneath sunspots and plages, extending to a depth of $\sim 1,000$ km. We are currently constrained by the low resolution of the observations from detailed mapping of the flow field and magnetic field profiles. It is tempting to believe that the sunspot which appears as a single large magnetic flux tube at the surface shreds into small flux fibres beneath this depth, and is held together as a coherent structure owing to the downflows¹. Moreover, downflows at a vertical flux tube result in inflows just beneath the surface (Fig. 2a) which could herd neighbouring flux tubes and form sunspots¹. This raises the question of what makes vertical magnetic flux tubes enter into either of these two distinct quasi-equilibrium states of sunspots and plages. Travel-time maps, coupled with theoretical calculations, might solve this puzzle when high-resolution observations become available. $\square$

Received 4 September; accepted 15 November 1995.

1. Parker, E. N. *Astrophys. J.* **230**, 905–913 (1979).
2. Duvall, T. L. Jr, Jefferies, S. M., Harvey, J. W. & Pomerantz, M. A. *Nature* **362**, 430–432 (1993).
3. Lee, W. H. K. & Peregry, V. in *Seismic Tomography, Theory and Practice* (eds Iyer, H. M. & Hirahara, K.) 9–22 (Chapman & Hall, New York, 1993).
4. Toner, C., Jefferies, S. M. & Duvall, T. L. Jr, *Astrophys. J.* (submitted).
5. Bracewell, R. N. *The Fourier Transform and its Applications* 2nd edn (McGraw-Hill, New York, 1986).
6. Bahcall, J. N. & Ulrich, R. K. *Rev. mod. Phys.* **60**, 297–372 (1988).
7. Woodard, M. F. & Libbrecht, K. G. *Astrophys. J.* **402**, L77–L80 (1993).
8. Kuhn, J. R. & Libbrecht, K. G. *Astrophys. J.* **381**, L35–L37 (1991).
9. Parker, E. N. *Astrophys. J.* **390**, 290–296 (1992).

ACKNOWLEDGEMENTS. The data in the figures in this Letter are available from the World Wide Web address <http://soi.stanford.edu/~duvall/duvall.html>. S.D. thanks J. Beckers, R. Howard and J. Harvey for encouragement and the opportunity to pursue independent research as a post-doctoral student. T.D. thanks P. Scherrer and the SOI group at Stanford for their hospitality during this work and for the use of their computing facilities, supported by NASA. This work was supported in part by the National Science Foundation, by the Solar Physics Branch of the Space Physics Division of NASA, the National Solar Observatory, and by the Office of Naval Research. The National Solar Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

Synthesis of RNA oligomers on heterogeneous templates

Gözen Ertem & James P. Ferris

Department of Chemistry, Rensselaer Polytechnic Institute, Troy, New York 12180, USA

THE concept of an RNA world^{1,2} in the chemical origin of life is appealing, as nucleic acids are capable of both information storage and acting as templates that catalyse the synthesis of complementary molecules³. Template-directed synthesis has been demonstrated for homogeneous oligonucleotides that, like natural nucleic acids, have 3',5' linkages between the nucleotide monomers⁴⁻⁷. But it seems likely that prebiotic routes to RNA-like molecules would have produced heterogeneous molecules with various kinds of phosphodiester linkages and both linear and cyclic nucleotide chains. Here we show that such heterogeneity need be no obstacle to the templating of complementary molecules. Specifically, we show that heterogeneous oligocytidylates, formed by the montmorillonite clay-catalysed condensation of actuated monomers, can serve as templates for the synthesis of oligoguanylates. Furthermore, we show that oligocytidylates that are exclusively 2',5'-linked can also direct synthesis of oligoguanylates. Such heterogeneous templating reactions could have increased the diversity of the pool of protonucleic acids from which life ultimately emerged.

The postulate that minerals catalysed the formation of biopolymers on the primitive Earth is supported by the observation that a montmorillonite clay catalyses the polymerization of activated mononucleotides⁸⁻¹⁴. Oligomers formed by the montmorillonite-catalysed condensation of the 5'-phosphorimidazole of cytidine (ImpC) Fig. 1a in aqueous solution (Fig. 2) contain 3',5'- and 2',5'-phosphodiester bonds, pyrophosphate (C⁵ppC) bonds, and linear and cyclic oligonucleotides (Tables 1, 2 and Fig. 1b). Template-directed synthesis^{15,16} was carried out using this heterogeneous mixture of oligo(C)s to determine whether complementary G oligomers would form. The hexamer and higher fractions of the C oligomers were separated from the shorter oligomers by Sephadex G-25 chromatography¹⁷ and by anion-exchange high-performance liquid chromatography (HPLC), and used as templates. The HPLC of the product mixture formed in the reaction

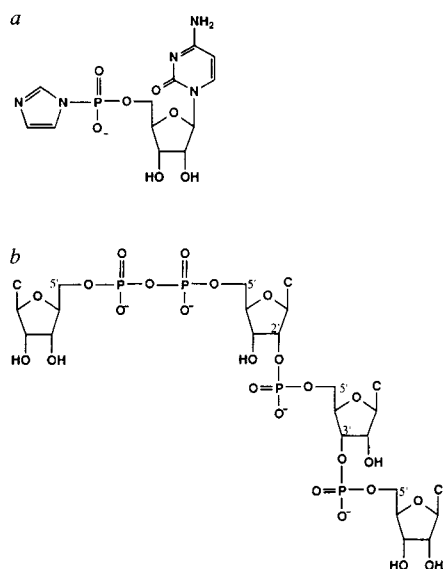


FIG. 1 The structures of a reactant and a product. a, The 5'-phosphorimidazole of cytidine (ImpC); b, C⁵ppC²pC³pC.

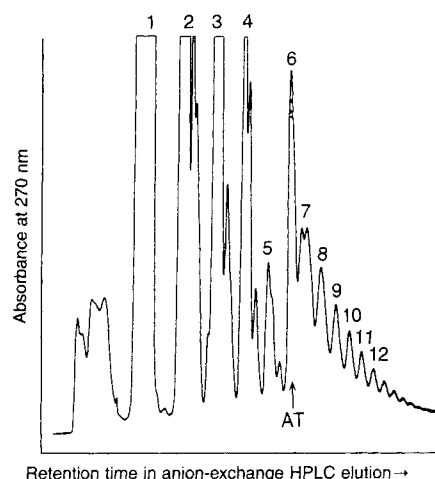


FIG. 2 Anion-exchange HPLC separation of the oligomers formed from the reaction of 1 ml 15 mM ImpC on 50 mg Na⁺-montmorillonite, prepared from Volelay, in 0.2 M NaCl and 0.075 M MgCl₂, pH 8, at room temperature for 5 days. (AT, change in attenuation). The presence of C⁵ppC in an oligomer has the same effect as a pC group on the retention times. Consequently, C⁵ppC is considered to be a monomer unit. Cyclic oligonucleotides have one less negative charge than the corresponding linear oligonucleotide. Therefore, they elute in the same HPLC fraction with the linear oligonucleotides containing one less monomer unit.

of 2-MeImpG with the oligo(C)s containing hexamers and higher oligomers revealed the formation of 3–6-mers of oligo(G)s (Fig. 3). Three peaks were observed for each of the trimer to hexamer fractions. Two control reactions were performed using the same reaction conditions. In the first, the condensation of ImpC was carried out in the absence of template and the longest oligomer formed was a trimer. This established that the tetramers to hexamers were formed by template reactions. In the second control, a previously reported template-directed synthesis using 3',5'-(Cp)₆₋₈ templates⁷ was reproduced. The detection in this control of mainly single peaks, corresponding to 3',5'-linked oligo(G)s, confirms that the complex product mixture shown in Fig. 2 is the result of template-directed synthesis on heterogeneous templates.

The triple peaks observed for each of the 3–6-mer oligo(G)s formed by template directed synthesis using oligomers obtained by the montmorillonite-catalysed reaction of ImpC (Fig. 3) suggest that oligomers containing both 2',5' and 3',5' linkages were formed. The observation of multiple peaks for each oligomer also suggests that 2',5'-linked oligo(C)s did not inhibit the oligo(G) synthesis but in fact served as templates, as only single peaks were

TABLE 1 Composition of oligo(C)s

Oligomer fraction	Linear oligomers (%)	Cyclic and CppC-containing oligomers (%)
'Monomers'	67	33
'Dimers'	85	15
'Trimers'	78	22
'Tetramers'	74	26
'Pentamers'	74	26

The composition of the monomer to pentamer fractions of the oligocytidylates was determined by collection of each fraction from a HEMA-IEC BIO-1000 Q column (Alltech)^{12,28}. Fractions (~1.5 mL) collected from the column were hydrolysed with alkaline phosphatase and the hydrolysis products separated by chromatography on the HEMA-IEC column.

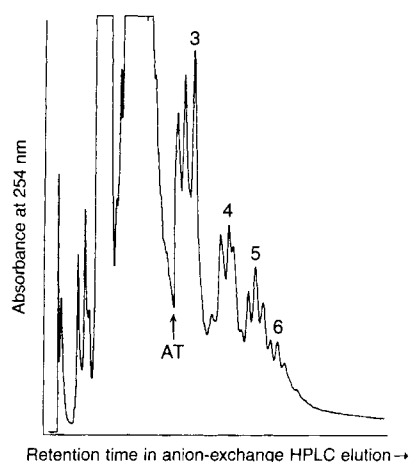


FIG. 3 Template-directed synthesis was performed in 10 μ l of solution containing 0.025 M oligo(C)s, 0.1 M 2-MelmpG, 1 M NaCl, 0.2 M MgCl₂, and 0.2 M HEPES, pH 8, for 6 days at 2 °C. (AT, change in attenuation). The oligo(G) chain lengths shown were established by comparison of their HPLC retention times with those of authentic samples of 3',5'-linked oligo(G)s prepared by the hydrolysis of poly(G). The Sephadex G-25 (DNA grade) column used in this study was prepared by Kamaluddin following the published procedure¹⁷. The oligo(C)s used as templates have the same retention times as the product oligo(G) dimers (Fig. 4).

TABLE 2 Structures of oligo(C)s formed on clay

Separated fraction	Composition (%)	3',5' links (%)
'Monomers'		
Hydrolysed by APH		
pC	100	
Not hydrolysed by APH		15
C ⁵ ppC	2.4	
[pC ² pC ²]	76	
[pC ³ pC ²]	12	
[pC ³ pC ³]	8.5	
'Dimers'		
Hydrolysed by APH		26
pC ² pC	73	
pC ³ pC	26	
Not hydrolysed by APH		
C ⁵ ppC ³ pC		
C ⁵ ppC ² pC		
[pC ² pC ² pC ²]		
[pC ² pC ² pC ³]		
[pC ² pC ³ pC ³]		
'Trimers'		
Hydrolysed by APH		44
pC ² pC ² pC	30	
pC ² pC ³ pC	36	
pC ³ pC ² pC	22	
pC ³ pC ³ pC	18	
Not hydrolysed by APH		
C ⁵ ppC ³ pC ² pC		
C ⁵ ppC ² pC ² pC		
[pC ² pC ² pC ² pC ²]		
[pC ² pC ² pC ³ pC ³]		
[pC ² pC ³ pC ³ pC ³]		

Individual fractions of the oligocytidylates, isolated by anion-exchange HPLC¹⁵, were first treated with alkaline phosphatase (APH) to cleave the 5'-terminal phosphate groups from the linear oligomers that do not contain a C⁵ppC grouping. The hydrolysed products, in which the terminal phosphate groups had been cleaved, were then separated from the unhydrolysed cyclic and C⁵ppC-containing oligomers by anion-exchange HPLC (Table 1). Each separated fraction was further characterised by selective cleavage of the 3',5'-phosphodiester bonds by RNase T₂ followed by APH hydrolysis as described¹². The complex nature of the non-hydrolysed portion of the dimer and trimer fractions, which contain both cyclic- and C⁵ppC-containing oligomers, together with the absence of authentic samples, made it impossible to determine the exact composition of these fractions.

observed when isolated 3',5'-linked oligo(C)s were used as templates.

We also investigated whether templates containing only 2',5'-linked oligomers could, on their own, facilitate the formation of the complementary oligomers, because some chemical processes designed to study the formation of RNA oligomers under prebiotic conditions produce mainly 2',5' isomers^{18,19}. The 2',5'-linked template was prepared by cleaving the 3',5'-linked phosphodiester bonds in the oligo(C)s formed by montmorillonite catalysis using ribonuclease (RNase) T₂ and then using the intact 2',5'-linked oligomers for template-directed synthesis. First, the pentamer and higher oligomers were isolated from the product mixture formed by clay-mineral catalysis using Sephadex G-25 chromatography. Then, those oligomers containing 3',5' linkages were cleaved with RNase T₂ and the 2',5'-linked pentamers and higher oligomers were again separated by Sephadex G-25 chromatography and anion-exchange HPLC. The resulting mixture, which contained linear and cyclic oligomers, and C⁵ppC-containing oligomers, was used as a template for oligo(G) formation. Tetramers, pentamers, and traces of hexamers were formed (Fig. 4), showing that oligo(G)s are produced from 2',5'-linked oligo(C) templates.

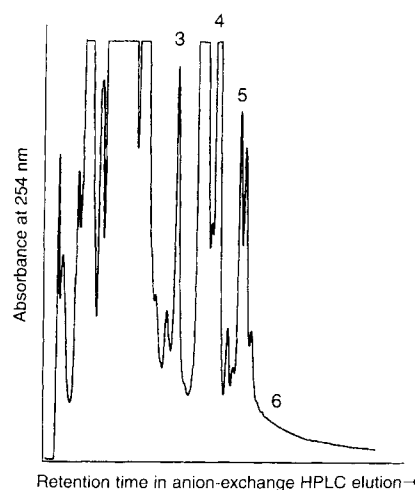


FIG. 4 Template-directed synthesis was performed in 7 μ l of solution containing 0.025 M template, 0.1 M 2-MelmpG, 1 M NaCl, 0.2 M MgCl₂ and 0.1 M HEPES, pH 8, for 19 days at 2 °C. Excess RNase T₂ was used to cleave the 3',5' linkages: control studies using 3',5'-linked oligo(C)s, obtained by the hydrolysis of poly(C), established that 50 units would have been required to cleave all the bonds in the synthetic oligo(C)s whereas 300 units were actually used in our experiments. An identical HPLC trace was obtained when this template-directed synthesis proceeded for 6 days.

The template peaks elute in the dimer region of the product oligo(G)s (not shown). Trimers were the longest oligomers formed when the templates were omitted from the reaction solution. The multiplicity of peaks observed, which had the HPLC retention times of tetramers and pentamers, suggests that the oligo(G)s formed have both 2',5' and 3',5' links. The synthesis of the complementary 2',5'-linked oligomers from a 3',5'-linked template has been reported²⁰. The converse process would proceed through a similar intermediate which contains hydrogen-bonded 2',5'- and 3',5'-linked oligomers²¹⁻²³.

It can be concluded from our studies that heterogeneous oligo(C)s formed by montmorillonite catalysis, or by almost any other chemical process²⁴, could have initiated the formation of complementary oligomers on the primitive Earth. The absence of regiospecificity in template-directed synthesis raises the question of how ribozymes with defined structures formed on the primitive Earth. One possibility is that catalysis of template-directed synthesis by minerals or metal ions resulted in greater regiospecificity⁸. In addition, some of the oligomers in the heterogeneous mixture formed by prebiotic processes could have catalysed the process of template-directed synthesis. This autocatalytic process would have selectively amplified the amounts of this catalytic RNA. Efficient ribozyme catalysts, which presumably had defined sequences, evolved from these initial transcripts²⁵⁻²⁷. □

Received 17 July; accepted 6 December 1995.

1. Joyce, G. RNA evolution and the origins of life. *Nature* **338**, 217-224 (1989).
2. *The RNA World* (eds Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Cold Spring Harbor, 1993).
3. Gilbert, W. *Nature* **319**, 618 (1986).
4. Inoue, T. & Orgel, L. E. *Science* **219**, 859-862 (1983).
5. Sievers, D. & von Kiedrowski, G. *Nature* **369**, 221-224 (1994).
6. Li, T. & Nicolaou, K. C. *Nature* **369**, 218-221 (1994).
7. Grzeskowiak, K. & Orgel, L. E. *J. molec. Evol.* **23**, 287-289 (1986).
8. Ferris, J. P. *Orig. Life Evol. Biophys.* **23**, 307-315 (1993).
9. Ferris, J. P. & Ertem, G. *Science* **257**, 1387-1389 (1992).
10. Ferris, J. P. & Ertem, G. *Orig. Life Evol. Biophys.* **22**, 369-381 (1992).
11. Ferris, J. P. & Ertem, G. *Orig. Life Evol. Biophys.* **23**, 229-241 (1993).
12. Ferris, J. P. & Ertem, G. *J. Am. chem. Soc.* **115**, 12270-12275 (1993).
13. Kawamura, K. & Ferris, J. P. *J. Am. chem. Soc.* **116**, 7564-7572 (1994).
14. Prabakar, K. J., Cole, T. D. & Ferris, J. P. *J. Am. chem. Soc.* **116**, 10914-10920 (1994).
15. Inoue, T. & Orgel, L. E. *J. Am. chem. Soc.* **103**, 7666-7667 (1981).
16. Orgel, L. E. *Nature* **358**, 203-209 (1992).
17. Hohn, T. H. & Schaller, H. *Biochem. biophys. Acta* **138**, 466-473 (1967).
18. Sawai, H. *J. Am. chem. Soc.* **99**, 7037-7039 (1976).
19. Sawai, H., Higa, H. & Kuroda, K. *J. chem. Soc. Perkins Trans. I* 505-508 (1992).
20. Lohrmann, R., Bridson, P. K. & Orgel, L. E. *Science* **208**, 1464-1465 (1980).
21. Kierzek, R., He, L. & Turner, D. H. *Nucleic Acids Res.* **20**, 1685-1690 (1992).
22. Jung, K.-E. & Switzer, C. J. *J. Am. chem. Soc.* **116**, 6059-6061 (1994).
23. Dougherty, J. P., Rizzo, C. J. & Breslow, R. *J. Am. chem. Soc.* **114**, 6254-6255 (1992).
24. Lohrmann, R. *J. Molec. Evol.* **18**, 185-195 (1982).
25. Doudna, J. A., Usman, N. & Szostak, J. W. *Biochemistry* **32**, 2111-2115 (1993).
26. Doudna, J. A., Couture, S. & Szostak, J. W. *Science* **251**, 1605-1608 (1991).
27. Green, R. & Szostak, J. W. *Science* **258**, 1910-1915 (1992).
28. Stribling, R. J. *Chromatography* **538**, 474-479 (1991).

ACKNOWLEDGEMENTS. Volclay, the mineral from which the sodium montmorillonite used in this study was prepared, was a gift from the American Colloid Company. The HPLC equipment used in this work was provided by NASA. This work was supported by NASA.

Economic and environmental choices in the stabilization of atmospheric CO₂ concentrations

T.M.L. Wigley*, R. Richels† & J.A. Edmonds‡

* University Corporation for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307-3000, USA

† Electric Power Research Institute, PO Box 10412, Palo Alto, California 94303, USA

‡ Pacific Northwest Laboratory, 901 D Street, SW, Suite 900, Washington DC 20024-2115, USA

THE ultimate goal of the UN Framework Convention on Climate Change is to achieve "stabilization of greenhouse-gas concentrations... at a level that would prevent dangerous anthropogenic interference with the climate system". With the concentration targets yet to be determined, Working Group I of the Intergovernmental Panel on Climate Change developed a set of illustrative pathways for stabilizing the atmospheric CO₂ concentration at 350, 450, 550, 650 and 750 p.p.m.v. over the next few hundred years^{1,2}. But no attempt was made to determine whether the implied emissions might constitute a realistic transition away from the current heavy dependence on fossil fuels. Here we devise new stabilization profiles that explicitly (albeit qualitatively) incorporate considerations of the global economic system, estimate the corresponding anthropogenic emissions requirements, and assess the significance of the profiles in terms of global-mean temperature and sea level changes. Our findings raise a number of important issues for those engaged in climate-change policy making, particularly with regard to the optimal timing of mitigation measures.

The IPCC Working Group I (WGI) concentration profiles (S350-S750; Fig. 1) were constructed under the following constraints: (1) prescribed initial (1990) concentration and rate of change of concentration; (2) a range of prescribed stabilization levels and attainment dates; and (3) the requirement that the implied emissions should not change too abruptly. Inverse calculations

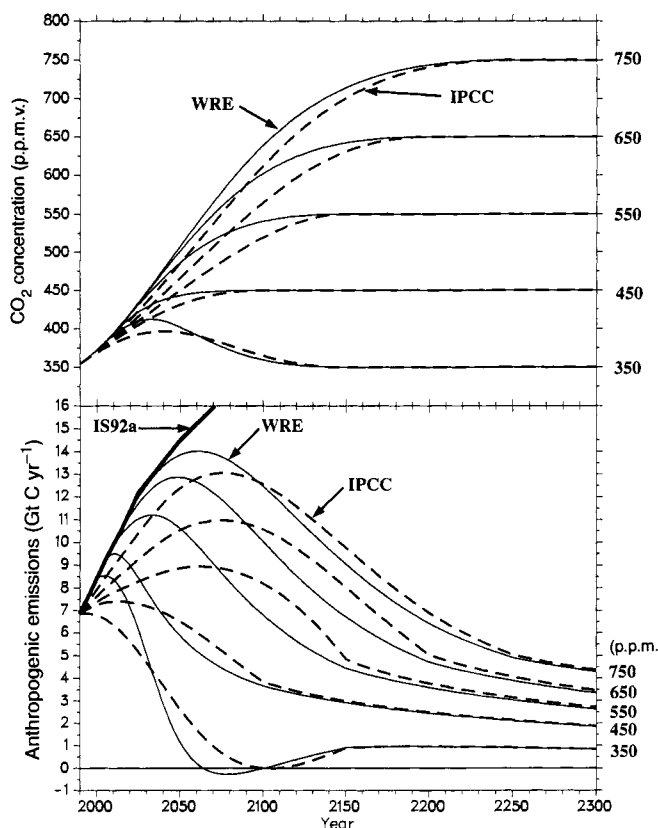


FIG. 1 Top, IPCC WGI^{1,2} (dashed lines) and revised concentration profiles (WRE (this paper), solid lines) for stabilization of CO₂ at 350-750 p.p.m.v. Bottom, implied anthropogenic emissions using the model of Wigley⁵. IS92a is shown (thicker line) for comparison. Emissions were calculated following the procedure in ref. 1 in which the terrestrial biosphere sink is characterized solely by CO₂ fertilization of net primary productivity. The implications of using CO₂ fertilization as the sole terrestrial sink are discussed in ref. 4. The post-1990 inverse calculations were initialized by specifying a value for the 1980s-mean net deforestation (D_{80s}). This determines the magnitude of the CO₂ fertilization factor. In the calculations in refs 1 and 2, D_{80s} was taken as 1.6 Gt C yr⁻¹. This value has subsequently been revised downwards to 1.1 Gt C yr⁻¹ (ref. 2), the value used here. Other minor budget changes have been made to accord with most recent data.

were then used to determine the emission rates required to achieve stabilization via the specified pathways. These show that stabilization requires an eventual and sustained reduction of emissions to substantially below current levels. Furthermore, some have interpreted the results for the IPCC pathways to imply that an immediate reduction in emissions (relative to the central IPCC "existing policies" or "business as usual" emissions scenario, IS92a³) is required to achieve any of the stabilization targets.

The WGI analysis was not intended as a recommendation for policy, but it will be carefully scrutinized for its policy implications. Consequently, it is important to understand what the analysis does and does not tell us. The first conclusion of the IPCC analysis, that meeting any of the prescribed targets will require emissions to decline eventually to levels well below today's, is robust. One cannot conclude from the WGI results, however, that an immediate reduction in emissions is required if we are to stabilize concentrations at 750 p.p.m.v. or below. The WGI emissions results correspond to just one of a range of possible pathways toward a particular concentration target. Stabilization at the same level, via different concentration routes, would produce different emissions.

What therefore are appropriate criteria for selecting a concentration (and hence emissions) time-path? Some guidance is found in the Framework Convention itself. Article 3 states that "policies and measures to deal with climate change should be cost-effective so as to ensure global benefits at the lowest possible cost". Thus, if two paths were indistinguishable in terms of their environmental

implications, then the path with the lower mitigation (that is, emissions reduction) costs would be preferred. If two paths differed in terms of their environmental impacts, the issue becomes one of balancing benefits and costs. Here we examine alternative pathways for meeting the prescribed concentration targets. We then consider both the economic (costs) and environmental (benefits) implications of choosing one concentration trajectory over another.

In revising the IPCC WGI profiles, we add an additional constraint to the three noted above: that the resulting emissions trajectories initially track a 'business as usual' (BAU) path. This is an idealization of the assumption that the initial departure from BAU would be slow. We also assume that the higher the concentration target, the longer the adherence to BAU. This produces quite different concentration pathways, complementing the ones defined by IPCC WGI.

If we constrain emissions to follow BAU initially, the required concentration paths must depend on what we assume for this baseline scenario. We concentrate here on results for the central IPCC scenario (IS92a³). A higher baseline (such as IS92e or f) will lead to higher initial emissions. For a lower baseline such as IS92c, the task of stabilization of CO₂ concentrations at a level around 500 p.p.m.v. would require little action⁴. To derive the new profiles, we followed IS92a concentrations for 10–30 years and then fitted a smooth curve to the stabilization levels and dates used in ref. 1, using the same Padé approximant method. Figure 1 compares the new profiles with the WGI profiles. Further details are given in ref. 4.

The emissions implied by these new pathways (Fig. 1, lower panel) were obtained using the model of Wigley⁵. Although the precise values are model-specific, as shown by the inter-model comparison of Enting *et al.*¹, the qualitative character of the results and the relative differences in emissions due to concentration pathway differences are not.

The WGI analysis suggests that an immediate departure from the BAU path is required to meet all CO₂ concentration targets. Figure 1 shows that this is not so for concentration targets of 450 p.p.m.v. and above. Furthermore, for targets of 550 p.p.m.v. and above, the maximum rates of emissions decline are similar in both the new and WGI cases (but more prolonged in the former).

Figure 2 compares the old (WGI) and new results in more detail for the 550 p.p.m.v. stabilization case, and assesses the sensitivity of the results to the length of the interval over which BAU emissions are followed. The upper panel shows the WGI pathway and revised pathways following BAU for 10, 20 and 30 years (the 20-year case is that considered earlier). The emissions differences (middle panel) are striking in terms of the implied carbon intensity of the global energy system in the early decades of the next century. The different cumulative emissions pathways diverge initially and then become nearly parallel as one approaches and moves beyond the stabilization point (AD 2150 in this case). Cumulative emissions are noticeably higher in the cases that follow IS92a initially (a result that applies to all stabilization levels). This is because the products of early emissions have a longer time to be removed from the atmosphere, and because the associated higher concentrations give stronger oceanic and terrestrial sinks. Thus, later emissions reductions allow greater total CO₂ production, particularly for higher stabilization levels. These cumulative emissions differences, not considered by IPCC², may have important economic implications.

We now turn to how mitigation costs might vary with the choice of concentration profile. The rising emissions baseline that we use corresponds to an assumption that, in the absence of policy intervention, CO₂ emissions will continue to grow. This is consistent with the overwhelming majority of studies recently reviewed by the IPCC⁶. The implication is that stabilizing concentrations will entail some positive mitigation costs. A growing baseline, however, does not imply the absence of "no regrets" emissions reduction options (that is, with zero or negative mitigation costs). Such options are typically included in sizeable

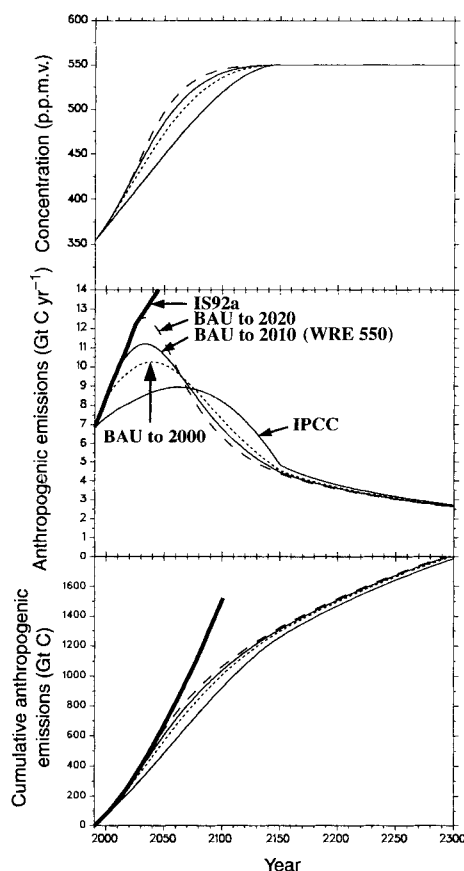
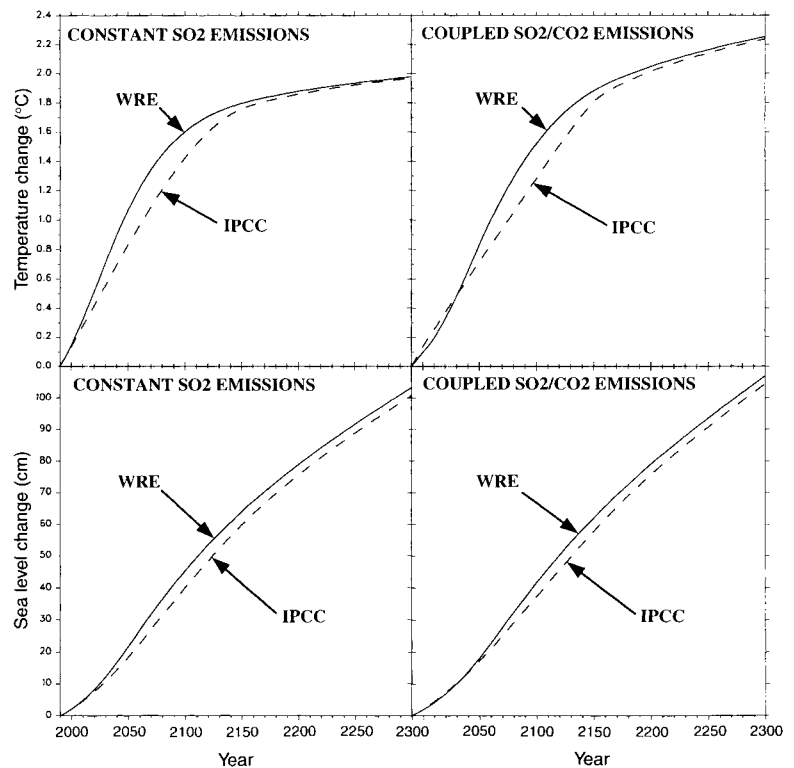


FIG. 2 Comparison of different concentration pathways (top panel) and implied emissions (middle) for stabilization of CO₂ levels at 550 p.p.m.v. in AD 2150. The pathways are: the original IPCC WGI S550 case¹; the revised profile shown in Fig. 1 based on following BAU background emissions for 20 years, from 1990 to 2010 (WRE 550); and alternative revised profiles in which BAU is followed by 10 (BAU to 2000) or 30 years (BAU to 2020). The bottom panel shows the corresponding cumulative emissions. IS92a values are shown for comparison (thicker dashed lines).

FIG. 3 Global-mean temperature (upper panels) and sea level changes (lower panels) for the S550 and WRE550 concentration stabilization pathways shown in Fig. 1. Results are from the models used in ref. 17, using the latest IPCC WGI estimate of the radiative forcing to 1990 and best-guess values of the climate sensitivity and ice-melt model parameters. Sea level changes include the contributions from oceanic thermal expansion, and ice melt from the world's glaciers and small ice caps, Greenland, and Antarctica. For the left panels, SO_2 emissions are held constant at their 1990 level. For the right panels, the effects of changes in anthropogenic SO_2 emissions (S) are added, with these changes directly coupled to those of fossil CO_2 emissions (F ; values from Fig. 1) using $S = [S(1990)/F(1990)]F$. In both cases, the effects of non- CO_2 greenhouse-gases are accounted for by scaling CO_2 forcing by 1.33, the mean scaling for the IS92 emissions scenarios (compare ref. 21).



quantities in most economic analyses^{7,8}. A growing baseline only means that economically competitive low-carbon alternatives are in insufficient supply to arrest future growth in carbon emissions. Conversely, if one were to assume that there are ample no-cost options to produce a falling emissions baseline, stabilization would entail little if any mitigation costs⁹.

Several analysts have studied how mitigation costs might vary with the timing of the emissions reductions. For example, Nordhaus¹⁰ and Manne and Richels¹¹ have identified cost-effective mitigation strategies for meeting a range of concentration targets. These studies show that to maintain cost-effectiveness, emissions tend to adhere longer to BAU the higher the concentration target (as assumed *a priori* here). Richels and Edmonds¹², in examining alternative emissions reduction pathways for stabilization at 500 p.p.m.v., found that the pathway can be just as important as the concentration stabilization level in determining the ultimate cost. Pathways involving modest reductions below a BAU scenario in the early years followed by sharper reductions later on were found to be less expensive than those involving substantial reductions in the short term. A similar conclusion can be found in Kosobud *et al.*¹³.

Viewing the stabilization issue as a carbon budget allocation problem helps explain why concentration pathways with higher near-term emissions have lower overall mitigation costs. Because cumulative emissions are approximately independent of the concentration pathway, for each stabilization level there is, roughly, a fixed allowable amount of CO_2 to be released. The basic choice is, therefore, how this budget is to be allocated over time. From this perspective, the reasons for drawing more heavily on the budget in the early years are: (1) *Positive marginal productivity of capital*. With the economy yielding a positive return on capital¹⁴, the further in the future an economic burden (here, emissions reduction) lies, the smaller is the set of resources that must be set aside today to finance the burden. (2) *Capital stock*. Stock for energy production and use is typically long-lived (for example, power plant, housing and transport). The current system is configured based upon a set of expectations about the future. Unanticipated changes will be costly. Time is therefore needed to reoptimize the capital stock. (3) *Technical progress*. There is ample evidence for past and potential future improvements in the efficiency of energy

supply, transformation and end-use technologies. Thus, the availability of low-carbon substitutes will probably improve and their costs reduce over time. In addition, as the emissions budget will be somewhat larger (that is, greater cumulative emissions) for pathways with higher emissions earlier, dependence on higher-cost, carbon-free alternatives is reduced.

We must stress that, even from the narrow perspective of a cost-effectiveness analysis, our results should not be interpreted as suggesting a "do nothing" or "wait and see" policy. First, all stabilization pathways still require future capital stock to be less carbon intensive than under a BAU scenario. As most energy production and use technologies are long-lived, this has implications for current investment decisions. Second, new supply options typically take many years to enter the market place. To ensure sufficient quantities of low-cost, low-carbon substitutes in the future requires a sustained commitment to research, development and demonstration today. Third, any available "no regrets" measures for reducing emissions should be adopted immediately. Last, it is clear from Fig. 1 that one cannot go on deferring emissions reductions indefinitely, and that the need for substantial reductions of emissions is sooner the lower the concentration target.

It is, of course, also important to examine the environmental consequences of selecting one concentration or emissions trajectory over another. This is because different concentration pathways imply, not only different emissions reduction costs, but also different benefits in terms of averted environmental impacts. In benefit-cost analyses of climate change policy options, it is common to use global-mean temperature and sea level rise as coarse indicators of the extent of climate impacts^{14,15}. We therefore calculate how these indicators are affected by differences in the pathways to stabilization at an atmospheric CO_2 concentration of 550 p.p.m.v., based on the model of Wigley and Raper^{16,17}. We first consider the direct effects of greenhouse-gas concentration changes, and then how these results may be modified by SO_2 emissions. All results use the central IPCC-recommended estimate of climate sensitivity¹⁸ (2.5°C equilibrium global-mean warming for a doubling of atmospheric CO_2 levels) and best-guess ice-melt model parameters¹⁷.

Figure 3 (left panels) shows that, if greenhouse gases alone are considered, both temperature change and sea level rise would be

noticeably affected by the choice of pathway towards stabilization at 550 p.p.m.v. These results, however, depend critically on how SO_2 emissions are assumed to change in the future. For the greenhouse-gas-alone case, we have assumed these emissions to remain constant at their 1990 level. As an alternative, we also consider a case where SO_2 emissions are closely coupled to fossil-fuel-derived CO_2 emissions. This case is consistent with the IPCC (IS92) emissions scenarios, except for IS92d, out to at least 2050 (ref. 3). It could occur if developing countries were less successful than developed countries in decoupling SO_2 and CO_2 emissions. In global-mean terms, SO_2/CO_2 emissions coupling leads to compensation between the reduced warming from reduced CO_2 emissions and an increased warming due to reduced SO_2 emissions^{19,20}.

To demonstrate the significance of this link, we give a specific example. This example is not meant to provide quantitative information on environmental impacts (which, for climate change, cannot be achieved through global-mean temperature alone), but to draw attention to aerosol influences as a critical factor in assessing the benefit–cost balance. Figure 3 compares global-mean temperature and sea level results for constant SO_2 emissions (left panels) with those for directly coupled CO_2 and SO_2 emissions (right panels). With SO_2 coupling, the lower-emissions case (S550) actually has warmer temperatures out to around 2040. This is because the much shorter lifetime of aerosols leads to a more rapid radiative forcing response to SO_2 emissions changes than to CO_2 emissions changes, allowing the former to dominate initially (compare refs 19,20). For sea level, coupling has a similar but less marked effect.

The market (for example, agriculture, timber and fisheries) and non-market (for example, biodiversity, environmental quality and human health) implications of these results are unclear: do pathway-related differentials up to $\sim 0.2^\circ\text{C}$ in global-mean temperature and 4 cm in global-mean sea level change translate into significantly higher damages and, if so, are these large enough to offset the reduced cost of a more economical transition away from fossil fuels? The answer depends on the regional details associated with these changes, and the sensitivities of impact categories to changes in important climate variables. Both aspects are highly uncertain. Nevertheless, it is clear that the choice of emissions path requires the consideration of both costs and benefits. $\square$

Received 30 May; accepted 8 December 1995.

- Enting, I. G., Wigley, T. M. L. & Heimann, M. *Future Emissions and Concentrations of Carbon Dioxide: Key Ocean/Atmosphere/Land Analyses* (Division of Atmospheric Res., CSIRO, Australia, 1994).
- Schimmel, D. S. et al. in *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emissions Scenarios* (eds Houghton, J. T. et al.) 35–71 (Cambridge Univ. Press, 1995).
- Leggett, J. A., Pepper, W. J. & Swart, R. J. in *Climate Change, 1992. The Supplementary Report to the IPCC Scientific Assessment* (eds Houghton, J. T., Callander, B. A. & Varney, S. K.) 69–95 (Cambridge Univ. Press, 1992).
- Wigley, T. M. L. in *The Global Carbon Cycle* (eds Wigley, T. M. L. & Schimmel, D. S.) (Cambridge Univ. Press, in the press).
- Wigley, T. M. L. *Tellus* **45B**, 409–425 (1993).
- IPCC Working Group III in *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emissions Scenarios* (eds Houghton, J. T. et al.) 233–304 (Cambridge Univ. Press, 1995).
- Dean, A. & Hoeller, P. *Costs of Reducing CO_2 Emissions* (OECD Economic Studies No. 19, OECD, Paris, 1992).
- Energy Modeling Forum *Reducing Global Carbon Emissions—Costs and Policy Options* (EMF 12, Stanford Univ., Stanford, CA, 1993).
- Goldemberg, J., Johansson, T. B., Reddy, A. & Williams, R. *Energy for a Sustainable World* (Wiley-Eastern Report, New Delhi, India & World Resources Institute, Washington, DC, 1987).
- Nordhaus, W. D. *The Efficient Use of Energy Resources* (Yale Univ. Press, New Haven, CT, 1979).
- Manne, A. & Richels, R. *The Greenhouse Debate—Economic Efficiency, Burden Sharing and Hedging Strategies* (Working Paper, Stanford Univ., Stanford, CA, 1995); *Energy J.* (in the press).
- Richels, R. & Edmonds, J. A. in *Integrative Assessment of Mitigation, Impacts, and Adaptation to Climate Change* (eds Nakicenovic, N., Nordhaus, W. D., Richels, R. & Toth, F. L.) 341–352 (International Institute for Applied Systems Analysis, Laxenburg, Austria, 1994); *Energy Policy* (in the press).
- Kosobud, R., Daly, T., South, D. & Quinn, K. *Energy J.* **15**, 213–232 (1994).
- Nordhaus, W. D. *Managing the Global Commons: The Economics of Climate Change* (MIT Press, Cambridge, MA, 1994).
- Fankhauser, S. *Valuing Climate Change* (Earthscan, London, 1995).
- Wigley, T. M. L. & Raper, S. C. B. *Nature* **357**, 293–300 (1992).
- Raper, S. C. B., Wigley, T. M. L. & Warrick, R. A. in *Rising Sea Level and Subsiding Coastal Areas* (ed. Milliman, J. D.) (Kluwer Academic, Dordrecht, The Netherlands, in the press).
- Mitchell, J. F. B., Manabe, S., Tokioka, T. & Meleshko, V. in *Climate Change. The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 131–172 (Cambridge Univ. Press, 1990).
- Wigley, T. M. L. *Nature* **349**, 503–506 (1991).
- Edmonds, J. A., Wise, M. & MacCracken, C. *Advanced Energy Technologies and Climate*

Change: An Analysis Using the Global Change Assessment Model (GCAM) (PNL-9798, UC-402, Pacific Northwest Lab., Richland, WA, 1994).

21. Wigley, T. M. L. *Geophys. Res. Lett.* **22**, 45–48 (1995).

ACKNOWLEDGMENTS. We thank the following for useful discussions: B. Bolin, H. Dowlatabadi, M. Grubb, E. Haites, A. Manne, R. Moss, W. Nordhaus, L. Pitelka, S. Smith, J. Weyant and L. Williams. The views expressed here, however, are solely those of the authors. This work was supported by the US Department of Energy and the Electric Power Research Institute.

Brief interstadial events in the Santa Barbara basin, NE Pacific, during the past 60 kyr

Richard J. Behl* & James P. Kennett

Marine Science Institute and Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA

THE instability of the Northern Hemisphere glacial climate over the past 100 kyr has been revealed by at least 20 brief warm (interstadial) episodes, called Dansgaard–Oeschger events, recorded in Greenland ice cores^{1–3} and in North Atlantic sedimentary records^{4,5}. A few of these events have been recognized elsewhere^{6–8}. Here we describe a record of ocean oxygenation and circulation from the Santa Barbara basin in the northeast Pacific Ocean which correlates well with the Greenland ice-core records. We see 19 of the 20 Dansgaard–Oeschger events, in the form of laminated sediments deposited under anoxic conditions, and we can correlate at least 16 of these with the 17 ice-core interstadials of the past 60 kyr. Thus, these short-term events were not restricted to the North Atlantic region. The events had substantial ecological and oceanographic effects in the Santa Barbara basin, including changes in benthic faunal populations and in the age and composition of bottom waters. Similar ventilation changes have been seen in the Gulf of California^{9,10}, suggesting that these changes may have been widespread and synchronous along the northeast Pacific margin. These results suggest sensitivity of broad areas of the ocean–atmosphere–cryosphere system to short-term climate change.

It is now generally accepted that switches in Quaternary climate conditions occurred at much shorter timescales than the orbitally modulated, 20 k to 100-kyr Milankovich-band cycles of insolation^{1–3}. Changes in temperature on millennial timescales are asymmetric, with abrupt, decadal-timescale warming followed by gradual cooling¹. To preserve records of such short-term fluctuations in the marine environment requires special conditions. Late Quaternary sedimentation rates in the Santa Barbara basin ($> 120 \text{ cm kyr}^{-1}$) are > 10 times higher than those typical of the deep sea, and owing to its geographical and bathymetric configuration the benthic environment is highly sensitive to climate and oceanographic changes, thus making this sequence well suited to the recording of changes at ultra-high resolution. The Santa Barbara basin is a confined, silled basin in the inner Continental Borderland of southern California¹¹. The modern basin is $\sim 600 \text{ m}$ deep and opens over a deep ($\sim 475 \text{ m}$) sill to the Pacific and a shallower ($\sim 230 \text{ m}$) sill to the adjacent Santa Monica basin. It contains oxygen-depleted waters below sill-depth, derived from the upper part of intermediate waters and the oxygen-minimum zone off central California (with $0.4\text{--}0.7 \text{ ml l}^{-1} \text{ O}_2$ at 500 m depth¹²). This source water is further depleted of oxygen ($< 0.1 \text{ ml l}^{-1} \text{ O}_2$) by degradation of organic matter from surface productivity (ref. 12). Oxygen deficiency excludes benthic macrofauna and prevents bioturbation¹³, so that below $\sim 525 \text{ m}$ seasonal variations in sediment supply (siliciclastic versus planktonic biogenic components) are preserved in the mud as thin, millimetre-scale laminations or

* Present address: Department of Geological Sciences, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, California 90840-3902, USA.

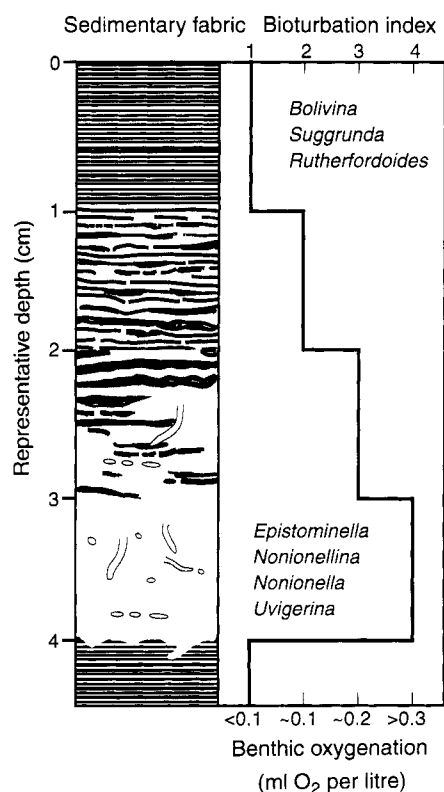


FIG. 1 Illustration of the bioturbation index used in this study. A value of one on this index represents unbioturbated sediments with distinct, continuous laminations (typically <1.0 to 2.0 mm thick); 2, diffuse, discontinuous or irregular laminations (1.5 to 4.0 mm thick), and rare 1–3-mm-thick horizontal *Chondrites* burrows; 3, slightly bioturbated with either faint, diffuse laminations or with few discrete patches of laminations surrounded by homogenized sediment, locally including 2–6-mm-diameter *Chondrites* burrows; 4, completely bioturbated fine-grained sediment displaying no discernible primary fabric and commonly containing macrofaunal echinoid spines. Characteristic benthic foraminiferal assemblages are listed for the two end-member facies. These bioturbation facies can be approximately correlated with oxygen concentration as follows: facies 1, <0.1 ml l⁻¹ O₂ (anaerobic conditions) that effectively exclude burrowing macrofauna; facies 2, 0.1 ml l⁻¹ O₂ (anaerobic/dysaerobic boundary conditions) allowing only micro- to meiofaunal bioturbation sufficient to diffuse but not destroy primary laminations; facies 3, 0.1–0.2 ml l⁻¹ O₂ (fluctuating anaerobic/dysaerobic boundary conditions) permitting partial vertical and lateral biological homogenization; facies 4, >0.3 ml l⁻¹ O₂ (dysaerobic to aerobic conditions) permitting the existence of bioturbating macrofauna (for example, echinoderms). The bioturbation index frequently changes within 1-cm increments and distinct burrows are never seen to penetrate vertically >2 cm; therefore, although major bioturbation events may modify several centimetres of previously deposited sediment, there is generally little correlation between the fabrics of consecutive centimetres of core.

'varves'. Documented subannual oxygenation events^{14,15} are insufficient to permit bioturbation, and laminations are well preserved in the recent sediment record^{16,17}.

Ocean Drilling Program (ODP) Hole 893A in the central Santa Barbara basin (at 34° 17.25' N, 120° 2.2' W; 576.5 m water depth) recovered 196.5 m of organic-rich silt and clay with sufficient abundance and preservation of planktonic and benthic foraminifera, diatoms and pollen to generate high quality palaeoclimate/palaeoceanographic records^{18–21}. Unlike the mostly laminated Holocene cores^{22,23}, Hole 893A penetrated an underlying sequence of alternately laminated and bioturbated (massive) hemipelagic muds, infrequently punctuated with fine- or coarse-grained event deposits^{18,24}.

The benthic foraminiferal oxygen-isotope record at Hole 893A correlates well with the standard deep-sea isotope stratigraphy of the latest Quaternary (SPECMAP)²⁵ suggesting continuous influence of the open ocean in Santa Barbara basin between oxygen

isotope stage 6 (~160 kyr) and the present²⁶. The chronostratigraphy of the past 60 kyr (0–83 m below sea floor) at Site 893 is based on a fit of 17 AMS radiocarbon ages²⁷ between 28,896 calendar years BP and the present ($r^2 = 0.998$), then on linear interpolation between three additional data points (3.13, 3.30 and 4.0) correlated with the SPECMAP^{25,28} marine oxygen-isotope stratigraphy between 30 and 60 kyr BP. The higher resolution and greater isotopic character of our record relative to the standard²⁶ prevents unequivocal assignment of a fourth data point (3.31). All geologically 'instantaneous' deposits, such as turbidites or storm layers > 10 cm thick, were removed from the linear extrapolation of the age–depth curve. No significant hiatuses are recognized and the sedimentation rate remains remarkably constant through glacial–interglacial eustatic sea level changes. The near-linear sedimentation rate may be due to the inverse influence of sea level on the basin's total depositional area and on the percentage of continentally derived sediment captured by submarine canyons and routed past Santa Barbara basin²⁴.

In Hole 893A, the degree of sediment bioturbation varies stratigraphically over scales of < 1 cm to > 10 m, corresponding to <10 to >10,000 yr. To quantify these changes, sediment texture and fabric were described at <1-cm increments²⁴ and a bioturbation index was established as described in the Fig. 1 legend. In the present basin and continental borderland, the distribution of benthic macro- and microfaunas, the degree of bioturbation, the

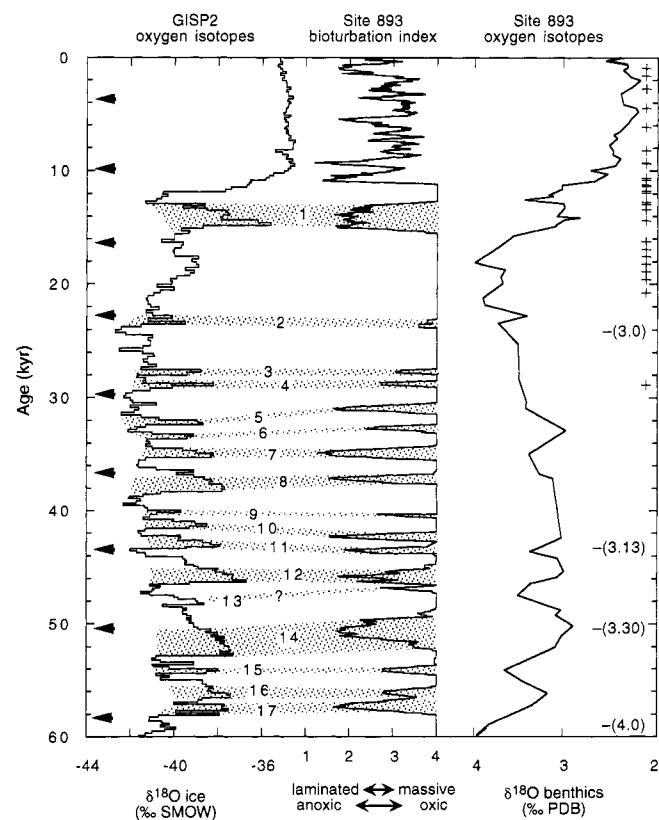


FIG. 2 Comparison of Site 893 bioturbation index and benthic foraminiferal $\delta^{18}\text{O}$ records with $\delta^{18}\text{O}_{\text{ice}}$ time-series from GISP2. Note the excellent correlation of Site 893 anoxia (lamination) events to 16 of 17 of the warm interstadials of GISP2. Bioturbation index is presented as a 49-cm (~300–400 yr) running average to dampen high-frequency variation and to match the resolution of the GISP2 record. Chronologies for GISP2 (ref. 7) and Site 893 were independently derived. Radiocarbon age control points (+) and SPECMAP data used for the Site 893 age model are shown to the right. Numbers in parentheses refer to standard data of the SPECMAP stratigraphy^{25,28}. The base of each core interval in Hole 893A is indicated by arrows to the left. Between 500 and 1250 yr may be missing at the two uppermost core breaks⁴³. Mean total error in analysis and calendar-year correction of the radiocarbon ages is 195 yr (ref. 27).

average depth and diameter of burrows, and the resulting sedimentary fabric are primarily controlled by the dissolved oxygen content of overlying bottom waters^{29,30}. Based on these data³⁰, the sediment facies in Hole 893A can be approximately correlated with oxygen concentration, as indicated in the figure legend.

We have also quantitatively analysed the composition of benthic foraminiferal assemblages associated with different bioturbation facies during the past 20 kyr. Two principal assemblages are apparent. Those associated with the Holocene laminated sediments are of low diversity and are dominated by varying abundances of low-oxygen-tolerant taxa *Bolivina*, *Suggrunda* and *Rutherfordoides*. Assemblages associated with non-laminated intervals of the last glacial episode exhibit higher diversity and are typical of well oxygenated bottom waters. These assemblages are dominated by *Epistominella* (~30–40%), *Nonionellina*, *Nonionella*, *Uvigerina* and miliolids such as *Pyrgo* and *Quinqueloculina*. There is an absence of *Suggrunda* and *Rutherfordoides* in these assemblages and *Bolivina* is rare to infrequent.

During the late Quaternary, basinal conditions switched between two end-member states: oxygen-deficient, with laminated sediments, absence of benthic macrofauna and dominance of benthic foraminifera tolerant of low oxygen levels; and oxygenated, with massive bioturbated sediments, presence of benthic macrofauna and dominance of foraminifera requiring higher oxygen levels. The Holocene (~11 kyr to present), the longest of the oxygen-deficient intervals, is predominantly laminated yet exhibits conspicuous variation in the degree of bioturbation (Fig. 2) that probably resulted from brief and subtle episodes of oxygenation. In spite of this variability, the complete dominance of low-oxygen benthic foraminiferal assemblages indicates that periods of oxygenation permitting bioturbation must have been short. Between 60 kyr and the base of the Holocene the bioturbation index indicates the presence of 16 principal anoxia events (laminated with low-oxygen foraminifera) separated by continuously oxygenated conditions (massive/bioturbated with higher-oxygen foraminifera) (Fig. 2). Individual anoxic episodes range from <200 to >2,500 yr (with a mean of 1,000 yr) with the longer events (12, 14, and the Holocene) displaying considerable centennial-scale variation²⁴; the other events are too short to capture much variability. High-frequency (~decadal) variability requires presentation of the data as a 49-cm (~300–400 yr) running average to best illustrate the major events recorded at Site 893. Although smoothing dampens expression of the rate of environmental change, the raw data indicate that most anoxic events began and ended abruptly, generally in <50–100 yr.

We can correlate the sequence from Hole 893A with the

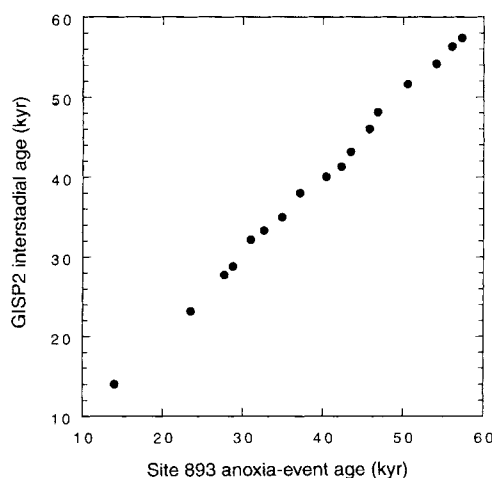


FIG. 3 Age correlation between Dansgaard-Oeschger interstadials exhibited by the $\delta^{18}\text{O}_{\text{ice}}$ record at GISP2 and anoxia events recorded at ODP Site 893, Santa Barbara basin; 13 of 16 events match to within 1% of the age of the interstadials; the others match to within 4%.

Greenland ice core (GISP2) over the past 60 kyr. The chronology for GISP2 was independently determined by Bender *et al.*⁷ based on a combination of annual layer counts³ and correlation with the deep-sea oxygen isotope record (SPECMAP)²⁵. A striking resemblance exists between the distribution, durations and magnitudes of palaeo-oxygenation events (bioturbation index) in Hole 893A and the Dansgaard-Oeschger interstadials recorded^{1–3} by oxygen isotopic changes ($\delta^{18}\text{O}_{\text{ice}}$) in GISP2 (Fig. 2). We identify anoxia events to match 19 of 20 interstadial events during the past 75 kyr and can tightly correlate with 16 of the last 17 interstadials identified in GISP2 since 60 kyr ago. Thirteen of these 16 episodes correlate to within 1% of their age (Fig. 3). A tight correlation is shown for all events younger than 30 kyr, which is within the most accurately dated interval using radiocarbon and annual layer counts. Tight correlations (no measurable offsets) also exist for events 7, 8, 12, and 15–17 (Fig. 2). Slight offsets in age are evident for 5 of 17 events. Interstadials 5 and 6 lead anoxia events by 900 yr (mean) and interstadials 9–11 lag anoxia events by 570 yr (mean). Systematic offsets for these groups of events, which are older than 30 kyr, suggest imperfections in the two SPECMAP-derived time-scales. The stratigraphic sequence of anoxic events >60 kyr ago also appears to match interstadials 18–20, but offsets are greater (2.5–5 kyr) than we can confidently correlate (probably due to unresolved problems with one or both chronologies) and is therefore not considered here.

A close relation has been demonstrated between changes in surface water temperature and the palaeo-oxygenation of the Santa Barbara basin since 20 kyr ago; laminated intervals and warm-water planktonic foraminiferal assemblages occur during the Bølling/Allerød and the Holocene, and cold-water assemblages and non-laminated intervals occur during the Last Glacial Maximum and the Younger Dryas⁹. Here we extend the observed relationship between climate and palaeoceanography by 40 kyr, through oxygen isotope stages 2 and 3. Differences between radiocarbon ages of coexisting benthic and planktonic foraminifera also show that laminated episodes during the past 20 kyr were associated with markedly older bottom waters (mean age 470 yr) than those of the massive (oxygenated) intervals (mean age 100 yr)^{9,27}.

The regional extent of coeval palaeo-oxygen fluctuations along North America's western continental margin is yet to be determined. However, a correlative succession of laminated and massive sediment is recorded^{9,10} at Deep Sea Drilling Program (DSDP) Site 480 in the Guaymas basin, Gulf of California, for the past 20 kyr, which, like Santa Barbara basin, also contains ~19 laminated intervals through stage 3. Several laminated intervals have also been reported^{31,32} from the open slope of the central California margin dating to oxygen isotope stages 2 and 3. We suggest that restricted circulation, extended residence time and high biological productivity in the Santa Barbara basin amplified small changes in the dissolved oxygen content of intermediate waters in the north-east Pacific during the late Quaternary, thereby recording palaeo-oxygenation fluctuations that may have been insufficient to affect the benthic ecology elsewhere along the margin. Although evidence exists for reduced fertility during glacial episodes in the eastern tropical north Pacific³³, diatom³⁰, foraminiferal¹⁹ and organic carbon studies^{34,35} at Site 893 found no consistent variation in productivity between deposition of laminated and bioturbated intervals, as is also the case for the Gulf of California³⁶.

The close correspondence between changes in basinal oxygenation and water-mass age⁹ implies that the bottom waters of Santa Barbara basin reflect the source and age of adjacent intermediate waters at near sill-depth. Intermediate water in the northeast Pacific is a mixture of several sources, including intermediate waters formed in the tropical east Pacific³⁷ and the Sea of Okhotsk³⁸. The large radiocarbon age of water currently flowing over the sill into Santa Barbara basin³⁹ (~1,300 yr) attests to the incorporation of a component of older water, originally derived from outside the Pacific basin. Variation in the relative contribution of any of these sources can modify the composition of

intermediate waters in the northeast Pacific and the associated concentration of dissolved oxygen. The absence of significant lags between changes in the GISP2 climate record and shifts in the palaeo-oxygenation at Hole 893A suggest that ocean circulation is tightly linked with global climate changes through the atmosphere. We cannot distinguish whether the character of northeast Pacific intermediate water was principally controlled by variation in production of young, proximally derived intermediate waters in the north Pacific or in flux of older, distally derived waters entering the Pacific basin. Shifts in atmospheric circulation associated with Northern Hemisphere cooling can rapidly change the location and flux of intermediate waters produced in the north Pacific⁴⁰, consequently influencing the ventilation of the Pacific. Likewise, deep-water production in the north Atlantic and flux of intermediate and deep water into the Pacific Ocean are also tightly linked with climate⁴¹. Changes in deep- and intermediate-water fluxes can be rapidly communicated to distant parts of ocean basins by shifting the production or contribution to other water masses⁴². For example, modelled shifts in deep-water production in the north Atlantic create changes in the upper part of the water column of the north Pacific within only a few decades (J. McWilliams, personal communication). Cessation or reduction of the flux of older 'conveyor-belt' water into the Pacific would increase the relative contribution of well oxygenated, proximally derived waters to intermediate depths in the northeast Pacific, producing the inverse relationship between water-mass age and palaeo-oxygenation found at Site 893. □

Received 21 August; accepted 6 December 1995.

1. Dansgaard, W. *et al.* *Nature* **364**, 218–220 (1993).
2. GRIP Project Members *Nature* **364**, 203–207 (1993).
3. Alley, R. B. *et al.* *Nature* **362**, 527–529 (1993).
4. Keigwin, L. D. & Jones, G. A. *J. geophys. Res.* **99**, 12397–12410 (1994).
5. Bond, G. & Lotti, R. *Science* **267**, 1005–1010 (1995).

6. Thouveny, N. *et al.* *Nature* **371**, 503–506 (1994).
7. Bender, M. *et al.* *Nature* **372**, 663–666 (1994).
8. Phillips, F. M., Campbell, A. R., Smith, G. I. & Bischoff, J. L. *Geology* **22**, 1115–1118 (1994).
9. Kennett, J. P. & Ingram, B. L. *Nature* **377**, 510–514 (1995).
10. Keigwin, L. D. & Jones, G. A. *Paleoceanography* **5**, 1009–1023 (1990).
11. Emery, K. O. *The Sea Off California: A Modern Habitat of Petroleum* (Wiley, New York, 1960).
12. CalCOFI Cruises 9301 & 9310 Physical, Chemical, and Biological Data **93-26 & 94-14**, (Univ. California, Scripps Institn of Oceanography, La Jolla).
13. Emery, K. O. & Hülsemann, J. *Deep-Sea Res.* **8**, 165–180 (1962).
14. Sholkovitz, E. R. & Gieskes, J. M. *Limnol. Oceanogr.* **16**, 479–489 (1971).
15. Reimers, C. E., Lange, C. B., Tabak, M. & Bernhard, J. M. *Limnol. Oceanogr.* **37**, 1577–1585 (1990).
16. Soutar, A. & Crill, P. A. *Geol. Soc. Am. Bull.* **88**, 1161–1172 (1977).
17. Schimmelmann, A. *et al.* *Mar. Geol.* **106**, 279–299 (1992).
18. Kennett, J. P., Baldauf, J. G. *et al.*, *Init. Rep. ODP Leg 146*, Part 2 (1994).
19. Kennett, J. P. & Venz, K. *Proc. ODP Sci. Res.* **146**, Part 2, 281–294 (1995).
20. Hemphill-Haley, E. & Fourtanier, E. *Proc. ODP Sci. Res.* **146**, Part 2, 233–250 (1995).
21. Heusser, L. E. *Proc. ODP Sci. Res.* **146**, Part 2, 265–280 (1995).
22. Soutar, A. & Crill, P. A. *Geol. Soc. Am. Bull.* **88**, 1161–1172 (1977).
23. Pisias, N. G. *Quat. Res.* **10**, 366–384 (1978).
24. Behl, R. J. *Proc. ODP Sci. Res.* **146**, Part 2, 295–308 (1995).
25. Martinson, D. G. *et al.* *Quat. Res.* **27**, 1–29 (1987).
26. Kennett, J. P. *Proc. ODP Sci. Res.* **146**, Part 2, 3–18 (1995).
27. Ingram, B. L. & Kennett, J. P. *Proc. ODP Sci. Res.* **146**, Part 2, 19–30 (1995).
28. Prell, W. L. *et al.* *Paleoceanography* **1**, 137–162 (1986).
29. Pearson, T. H. & Rosenberg, R. *Oceanogr. mar. Biol. a. Rev.* **16**, 229–311 (1978).
30. Savrda, C. E., Bottjer, D. J. & Gorsline, D. S. *Bull. Am. Ass. Petrol. Geol.* **68**, 1179–1192 (1984).
31. Hemphill-Haley, E. & Gardner, J. V. *Quat. Res.* **41**, 131–135 (1994).
32. Gardner, J. V. & Dartnell, P. in *IGCP 374: Palaeoclimatology and Palaeoceanography from Laminated Sediments* (eds Behl, R. & Kennett, J.) 11 (Univ. California, Santa Barbara, 1995).
33. Ganeshram, R. S., Pedersen, T. F., Calvert, S. E. & Murray, J. W. *Nature* **376**, 755–758 (1995).
34. Gardner, J. V. & Dartnell, P. *Proc. ODP Sci. Res.* **146**, Part 2, 103–124 (1995).
35. Stein, R. & Rack, F. R. *Proc. ODP Sci. Res.* **146**, Part 2, 125–138 (1995).
36. Sancetta, C. *Paleoceanography* **10**, 67–84 (1995).
37. Reid, J. L. *Intermediate Waters of the Pacific Ocean* (Johns Hopkins Press, Baltimore, 1965).
38. Talley, L. D. *Deep-Sea Res.* **38**, S171–S190 (1991).
39. Ostlund, H. G. & Stuiver, M. *Radiocarbon* **22**, 25–53 (1980).
40. Keigwin, L. D., Jones, G. A. & Froelich, P. N. *Earth planet. Sci. Lett.* **111**, 425–440 (1992).
41. Broecker, W. S. & Denton, G. H. *Geochim. cosmochim. Acta* **53**, 2465–2501 (1989).
42. Toggweiler, J. R. & Samuels, B. *Deep-Sea Res.* **42**, 477–500 (1995).
43. Merrill, R. B. & Beck, J. W. *Proc. ODP Sci. Res.* **146**, Part 2, 45–59 (1995).

ACKNOWLEDGEMENTS. We thank the personnel of the Ocean Drilling Program for invaluable assistance, and T. Sowers and M. Bender for help with the GISP2 chronology. This work was supported by the JOI/USSP, NSF, DOE/NIGEC/WESTGEC and a fellowship (R.J.B.) from the UCSB Marine Science Institute.

Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700

Kenji Satake*, Kunihiro Shimazaki†, Yoshinobu Tsuji† & Kazue Ueda†

* Seismotectonics Section, Geological Survey of Japan, Tsukuba 305, Japan

† Earthquake Research Institute, University of Tokyo, Tokyo 113, Japan

Geological evidence shows that great earthquakes have occurred in the recent prehistoric past in the Cascadia subduction zone, off the Pacific coast of North America. The most recent event (or series of events) is dated at about 300 years ago^{1–4}, but the precise date and magnitude have not been determined. Geological investigations have not been able to distinguish a single giant earthquake from a series of great earthquakes occurring over a timespan of a decade or two⁴, although this information is important for the assessment of future hazard⁵. We have found several tsunami records in Japan from AD 1700 with no indication of a local cause. Historical earthquake records and palaeoseismic evidence indicate the absence of a large earthquake in 1700 in South America, Alaska or Kamchatka, leaving Cascadia as the most likely source of this tsunami. The estimated time of the earthquake is the evening (about 21:00 local time) of 26 January 1700. The magnitude is estimated as 9 from the tsunami heights, in which case the earthquake ruptured the entire length of the Cascadia subduction zone². These estimates are consistent with

Native American legends that an earthquake occurred on a winter night⁶.

No great (moment magnitude $M_w \approx 8$) or giant ($M_w \approx 9$) earthquake has been historically or instrumentally recorded in the Cascadia subduction zone, where the Juan de Fuca plate descends beneath the North American plate (Fig. 1)^{5,7}. By comparison, in many other subduction zones with a similar geological setting, such as southwestern Japan or southern Chile, occurrence of great subduction-zone earthquakes has been documented. For example, Japanese historical documents describe damage from great earthquakes and tsunamis along the undersea Nankai trough, southwest Japan, approximately every 100 years^{8,9}. However, the relatively short documentary history (starting in the late eighteenth century at the earliest) prevents the use of American written records to estimate the size and recurrence interval of earthquakes at the Cascadia subduction zone. Several oral traditions of natives on the coast from northern California to British Columbia describe an earthquake and tsunami of unknown date^{6,10–13}.

Geological evidence for subduction earthquakes in Cascadia has been accumulated in the past decade (Fig. 1)^{1–3} and radiocarbon and tree-ring dating of the most recent event indicates that it occurred within a decade or two of AD 1700^{3,4}. Evidence distributed from California to British Columbia indicates that the entire Cascadia subduction zone was ruptured either by one $M_w \approx 9$ earthquake or by a series of $M_w \approx 8$ events⁴. A comparison of various geological features of subduction zones⁵ suggests that the maximum size of Cascadia earthquakes is $M_w \approx 9.5$. A more recent study of geological features of the overlying plate¹⁴ suggests that the maximum size is $M_w \approx 8$.

The earthquake magnitude affects ground shaking¹⁵ and tsunami height¹⁶, and hence damage to humans and property. If a giant ($M_w \approx 9$) earthquake occurs in the future, the seismic and tsunami hazards in the Pacific Northwest would be considerable⁵ and the tsunami would probably cause damage across the Pacific

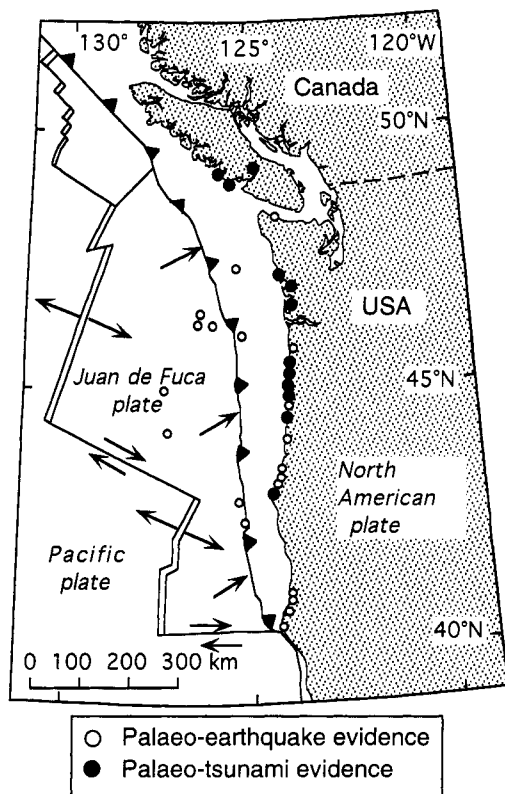


FIG. 1 Map of the Cascadia subduction zone and sites of palaeo-earthquake (open circles) and palaeo-tsunami (solid circles) evidence for the 300-year-old event(s)². The palaeo-earthquake evidence includes offshore turbidites, buried soils and dead forests that indicate sudden coastal subsidence, and liquefaction on land due to ground shaking. The palaeo-tsunami evidence is sand layers deposited by tsunamis. The arrows indicate relative movements of the plates at their boundaries.

Ocean including Hawaii and Japan. It is therefore important to estimate directly the size of the past earthquake(s).

In Japan, where historical documents of earthquakes have been collected and analysed, tsunami damage from an unknown source was documented at several places on 27–28 January 1700, by the gregorian calendar¹⁷. At Miyako on the Pacific coast (see Fig. 2), 20 houses were damaged. The tsunami height there is estimated as 2–3 m from our field measurements based on the description of historical documents, with an error no larger than 0.5 m. At Otsuchi, about 30 km south on the coast, another documents describes a tsunami arriving around midnight on 27 January, inundating rice paddies and damaging two houses. We estimate the tsunami height as 3 m. At Tanabe, approximately 1,000 km southwest along the coast, flooding of a storehouse floor the next morning (28 January) was reported, indicating a 2-m tsunami. In addition to these three localities, independently reported in four different documents, a tsunami of unknown height on that date was documented at two other locations, Nakaminato and Miho (Fig. 2).

These waves have features characteristic of a far-field tsunami. The wide and apparently uniform distribution of water heights (Fig. 2) excludes the possibility of a meteorological origin such as a storm surge. Most storm surges in Japan are caused by typhoons, which are observed only from August to October. In fact, the weather on 27–28 January 1700 was reported to be sunny or cloudy in central Japan. There is no record of an earthquake in Japan on that day, which excludes the possibility of a tsunami from a local earthquake. The Pacific coast of Japan has experienced tsunamis from far-field earthquakes around the Pacific Ocean several times in this century. For example, the tsunami from the 1960 Chilean earthquake (M_w 9.5) travelled across the Pacific Ocean in about 24 hours and caused extensive damage including about 140 fatalities in Japan¹⁸. The tsunamis from the 1964 Alaska (M_w 9.2) and 1952 Kamchatka (M_w 9.0) earthquakes were much smaller (Fig. 2) yet caused some damage. The smaller tsunamis were partly due to the directivity effect; tsunami heights are usually smaller in the direction parallel to the orientation of the fault¹⁹. Abe^{20,21} defined a tsunami magnitude scale, M_t , which represents an earthquake size estimated from observed tsunami heights and has been calibrated in such a way that it is approximately equal to the moment magnitude M_w .

We estimate earthquake magnitudes for possible source regions of the 1700 tsunami and examine historical and geological evidence for such an event in each region. The possible sources

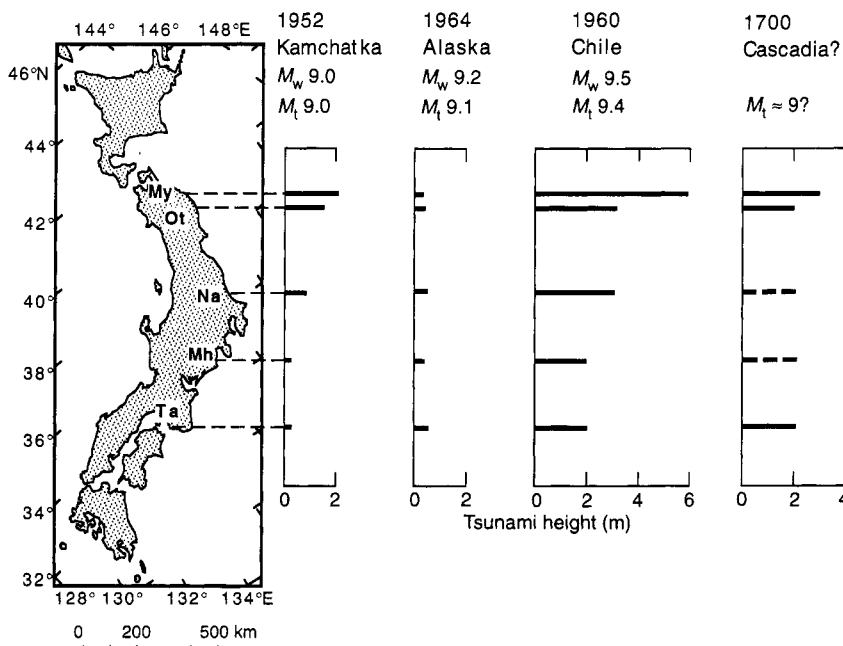


FIG. 2 Tsunami heights along the Japanese coast (at the locations where the AD 1700 tsunami was observed) from the 1952 Kamchatka, 1964 Alaska and 1960 Chile earthquakes, in addition to the 1700 tsunami from an unknown source¹⁸ (My, Miyako; Ot, Otsuchi; Na, Nakaminato; Mh, Miho; Ta, Tanabe). Earthquake moment magnitude M_w and tsunami magnitude M_t are also shown for each event. The dashed lines for the 1700 tsunami indicate that the tsunamis were observed, but the estimation of heights is difficult.

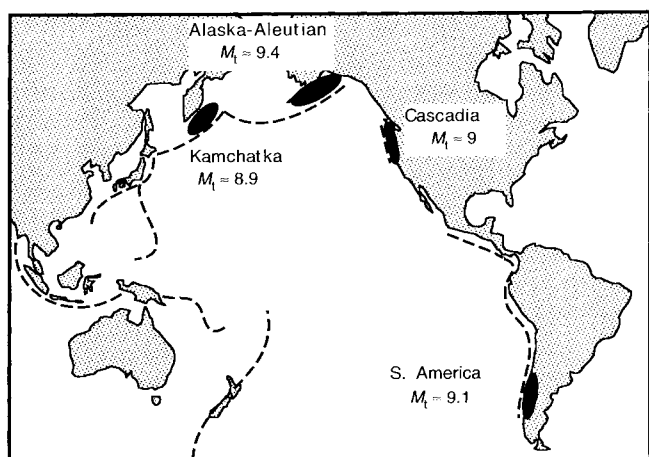


FIG. 3 Possible source regions for the 1700 tsunami observed in Japan and the estimated earthquake size M_t derived from the tsunami heights. We used $M_t = \log H_t + B$ where H_t is the local mean tsunami run-up (or inundation) height in metres, which has been shown to be equal to twice the peak-to-trough amplitude measured by tide gauges²¹. We used $H_t = 2$ m from our observations. The constant B is 8.8 for South America, 9.1 for Alaska-Aleutian and 8.6 for Kamchatka sources. Maximum tsunami run-up height on land has been shown to be equal to $\sim 2 H_t$ (refs 21, 29). If we use the maximum observed height (3 m at Otsuchi and Miyako) as the maximum run-up height, then the M_t estimates are 0.1 smaller than those shown here, which indicates a possible error range. For the Cascadia source, however, both the M_t and error estimates are difficult because of lack of knowledge of the appropriate value for B .

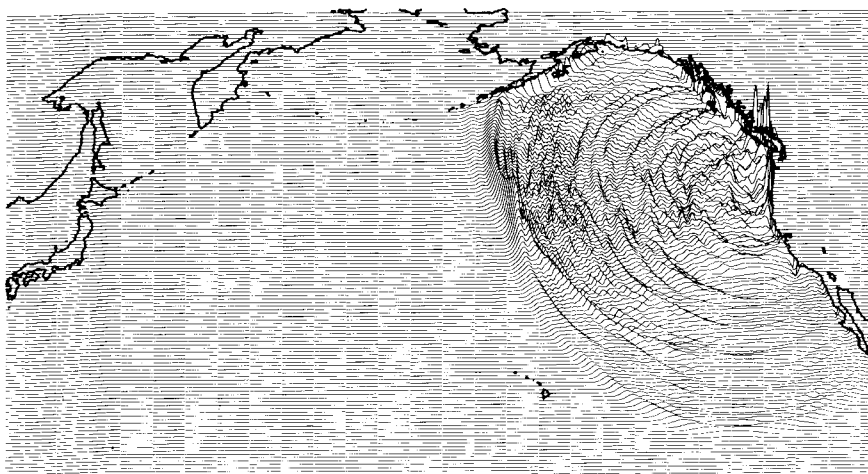
of the 1700 earthquake besides Cascadia are South America, Alaska-Aleutian and Kamchatka²⁰. If the tsunami came from South America, the earthquake would be $M_t = 9.0-9.1$; for the Alaska or Aleutian source $M_t = 9.3-9.4$; and for the Kamchatka source $M_t = 8.8-8.9$. Written documents of past earthquakes exist for South America around AD 1700 and such a great earthquake should have been documented. For example, the 1687 Peru (Callao) earthquake (M 7.6), the 1730 Valparaíso earthquake (M 8.7) and 1751 Concepción earthquake (M 8.5) generated tsunamis and were recorded in Japan^{18,22,23}. However, there is no corresponding earthquake record in South America for 26 January 1700. For the Kamchatka region, Russian settlers started to immigrate in the 1680s, and the oldest record of an earthquake is 1737²⁴. There is no written record or corresponding geological evidence for a large earthquake in 1700 (V. K. Gusiakov, personal communication). In addition, tsunami height from a Kamchatka source would decay with distance along the Japanese coast as demonstrated by the 1952 earthquake (see Fig. 2). Tsunami heights from earthquakes in Alaska or the Aleutian Islands are not very large in Japan because of the directivity effect. For example, they were mostly less than 0.3 m for 1964 Alaskan earthquake (M_w 9.2). Therefore, to produce the observed tsunami heights in Japan, the earthquake source must be larger. There has been no geological or historical evidence found in the region for such an event^{25,26}. On the other hand, palaeoseismological evidence for a great or giant earthquake exists in the Cascadia subduction zone, hence we conclude that Cascadia is the most likely source of the 1700 tsunami.

If the 1700 tsunami came from Cascadia, we can estimate the origin time and magnitude of the earthquake. The earliest documented tsunami arrival time was around midnight on 27 January

Japan time, or around 15:00 on 27 January GMT. Because tsunami travel time from Cascadia to Japan is about 10 hours, the earthquake origin time is estimated at around 5:00 on 27 January GMT or 21:00 on 26 January local time in Cascadia. This time is consistent with Native American legends that an earthquake occurred on a winter night⁶.

We estimate the moment magnitude of this earthquake to be 9. Although the tsunami magnitude formula²⁰ has not been established for a Cascadia earthquake owing to lack of instrumental data, the estimated magnitude values for all other possible regions are about 9 as we have already seen (Fig. 3). We can also simulate tsunami propagation across the Pacific Ocean by applying a finite-difference method to real bathymetry data²⁷. Preliminary computations indicate that a Cascadia earthquake with $M_w \approx 8$ produces only a 0.3-m tsunami in Japan, too small to cause damage and easily missed if it happened at night. If the earthquake size is $M_w \approx 9$ (Fig. 4), the tsunami height becomes 2 m, large enough to cause damage. A recent earthquake provides observational support in a reciprocal way; an earthquake ($M_w = 8.2$) occurred in the Kuril Islands, near Japan, on 4 October 1994²⁸, and the tsunami heights were only 0.3 m or less on the Pacific coast of California, Oregon and Washington. A Cascadia earthquake with $M_w \approx 8$ would therefore produce similar tsunami heights in Japan. (The rupture length of an $M_w \approx 8$ earthquake is only $\sim 100-200$ km, so that directivity effects are negligible.) An earthquake with $M_w \approx 9$ extending along the entire Cascadia zone, on the other hand, would cause a damaging tsunami in Japan, which lies in the direction perpendicular to the fault, because of the directivity effect as well as the larger deformation of the ocean bottom. □

FIG. 4 Snapshot of computed tsunami propagation across the Pacific Ocean from a Cascadia earthquake. The tsunami propagation is simulated using a finite-difference method²⁷ and digital bathymetry data of the area shown. The ocean-bottom deformation calculated from an M_w 9 earthquake extending along the entire Cascadia subduction zone is used as an initial condition. Water height distribution 4 hours after the earthquake is shown here. The peak water height in the central Pacific north of Hawaii is about 1 m. The large tsunami height towards Japan is due to the directivity effect; the tsunami amplitude is larger in a direction perpendicular to the fault orientation.



Received 3 October; accepted 4 December 1995.

- Atwater, B. F. *Science* **236**, 942–944 (1987).
- Atwater, B. F. *et al.* *Earthquake Spectra* **11**, 1–18 (1995).
- Atwater, B. F., Stuiver, M. & Yamaguchi, D. K. *Nature* **353**, 156–158 (1991).
- Nelson, A. R. *et al.* *Nature* **378**, 371–374 (1995).
- Heaton, T. H. & Hartzell, S. H. *Science* **236**, 162–168 (1987).
- Clague, J. J. *Current Research, Geological Survey of Canada* **1995-A** 47–50 (1995).
- Heaton, T. H. & Kanamori, H. *Bull. seism. Soc. Am.* **74**, 933–941 (1984).
- Ando, M. *Tectonophysics* **27**, 119–140 (1975).
- Sangawa, A. *Jishin-koukougaku (Seismoarchaeology)* (Chuou-kouron, Tokyo, 1992).
- Kreober, A. L. *Yurok Myths* (Univ. California Press, Berkeley 1976).
- Heaton, T. H. & Snavely, P. D. *Bull. seism. Soc. Am.* **75**, 1455–1460 (1985).
- DeBois, C. A. *Am. Anthropol.* **34**, 248–262 (1932).
- Maud, R. (ed.) *The Secheit and the South-eastern Tribes of Vancouver Island* (Talonbooks, Vancouver, 1978).
- MacCaffrey, R. & Goldfinger, C. *Science* **267**, 856–859 (1995).
- Cohee, B. P., Somerville, P. G. & Abrahamson, N. A. *Bull. seism. Soc. Am.* **81**, 28–56 (1991).
- Whitmore, P. M. *Natural Hazards* **8**, 59–73 (1993).
- Tsuji, Y. *Rekishi Jishin (Historical Earthquakes)* **3**, 220–238 (1987).
- Watanabe, H. *Nihon Higai Tsunami Souran (Catalogue of Damaging Tsunamis in Japan)* (Univ. Tokyo Press, 1985).
- Ben-Menahem, A. & Rosenman, M. J. *geophys. Res.* **77**, 3097–3128 (1972).
- Abe, K. J. *geophys. Res.* **84**, 1561–1568 (1979).
- Abe, K. *Bull. Earthquake Res. Inst. Univ. Tokyo* **64**, 51–69 (1989).
- Lockridge, P. A. *Tsunamis in Peru-Chile* (National Geophys. Data Center, Boulder, 1985).
- Lomnitz, C. *Geol. Rdsch.* **59**, 938–960 (1970).
- Soloviev, S. L. *Basic Data on Tsunamis on the Pacific Coast of the USSR, 1737–1796, Study of Tsunami in an Open Ocean* (Nauka, Moscow, 1978).
- Combellick, R. A. *Professional Rep. 112*, (Alaska Div. geol. geophys. Surveys, Fairbanks, 1991).
- Plafker, G., Lajoie, K. R. & Rubin, M. in *Radiocarbon After Four Decades* (eds. Taylor, R. E., Long, A. & Kra, R. S.) 436–453 (Springer, New York, 1992).
- Satake, K. *Pure appl. Geophys.* **144**, 455–470 (1995).
- Tanioka, Y., Ruff, L. & Satake, K. *Geophys. Res. Lett.* **22**, 1661–1664 (1995).
- Kajiura, K. in *Tsunamis—Their Science and Engineering* (eds. Iida, K. & Iwasaki, T.) 131–145 (Terra Sci. Pub., Tokyo, 1983).

ACKNOWLEDGEMENTS. This work was initiated from conversations at Paleoseismology Workshop in Marshall, California, in September 1994. We thank the workshop organizers (B. Yeats, D. Schwartz and C. Prentice), and A. Nelson for discussion. We also thank G. Carver and R. Hyndman for information on Native American legends, and K. Abe, B. Atwater and K. Okumura for discussions. We also thank C. Allen, B. Atwater, J. Bourgeois, A. Moore, A. Nelson and E. Okal for reviews. This work was partially supported by the Japanese Earthquake Prediction Program.

Manufacture and use of hook-tools by New Caledonian crows

Gavin R. Hunt

Department of Ecology, Massey University, Private Bag 11222, Palmerston North, New Zealand

TOOL behaviour in wild birds has been described as mostly stereotyped^{1,2}, and tool manufacture involves little modification of material^{3–5}. Here I report in New Caledonian crows *Corvus moneduloides* the manufacture and use of two different types of hook tool to aid prey capture: hooked-twig and stepped-cut barbed pandanus leaf. Crow tool manufacture had three features new to tool use in free-living nonhumans: a high degree of standardization, distinctly discrete tool types with definite imposition of form in tool shaping, and the use of hooks. These features only first appeared in the stone⁶ and bone⁷ tool-using cultures of early humans after the Lower Palaeolithic^{6,7}, which indicates that crows have achieved a considerable technical capability in their tool manufacture and use.

Crows are forest birds endemic to New Caledonia, where they are omnivores and live in family groups⁸. Like Corvids generally^{9–14} they are resourceful, for example manipulating twigs with their bills to search for prey in dead wood¹⁵ and dropping *Aleurites moluccana* nuts onto rocks to get their seeds. In rain-forest on Pic Ningua (950–1,300 m above sea level), I observed non-banded, non-aged and non-sexed crows between November 1992 and March 1995 in all months except January. On 52 different occasions between 07:00 and 15:30 I observed tool behaviour by one or more (up to four) birds. Observations were made of four crows manufacturing tools and 68 crows using or carrying tools. On average, three birds (range, 1–9) were present

TABLE 1 Data on stepped-cut tool cutouts from *Pandanus* spp. trees at two areas on Pic Ningua, and one at Mt. Cindoa.

Tree no.	No. cutouts collected per tree	No. cutouts estimated as present	Mean cutout length (cm) for each area/s.e.m./range	Mean no. of steps per cutout for each area/s.e.m./range
Pic Ningua (area 'A')				
1	13	17		
2	23	30		
3	18	21	14.95/0.172/	3.10/0.037/
4	37	53	(10.6–22.2 cm)	(2–5 steps)
5	14	15		
Pic Ningua (area 'B')				
6	17	18		
7	52	72		
8	23	28	16.65/0.170/	3.11/0.037/
9	51	65	(12.1–24.9 cm)	(2–6 steps)
10	14	14		
Mt Cindoa				
11	9	9		
12	10	10		
13	5	5	20.65/0.644/	3.49/0.102/
14	7	7	(11.8–40.0 cm)	(2–5 steps)
15	12	13		

Pandanus spp. trees can be single or multitrunked, with long, narrow leaves at the top of a trunk that spiral outwards and upwards around it. Trees 12 and 15 were double-trunked. I chose trees from two separate areas on Pic Ningua (minimum distance separating the trees between the two areas was 900 m) to try and detect differences in the structure of tools manufactured by crows. I chose a tree from a distance without knowledge of the number or shape of cutouts on it, and finally selected it if sufficient cutouts were present (> 9 at Pic Ningua; > 4 at Mt Cindoa). The five selected trees within each area were widely distributed. I removed as many undamaged cutouts as possible from living leaves (cutouts were mostly near leaf bases). I measured cutout length (length of missing leaf edge) off the (undried) cutout shape I outlined on paper. Mean cutout length was different between all three areas when the significant variation between trees within each area was removed (Nested ANOVA, $F = 21.88$, d.f. = 2, $P \ll 0.05$; and $F = 4.37$, d.f. = 12, $P \ll 0.05$, respectively). Available space (estimated number of cutouts present) on leaf edges of individual trees at Mt Cindoa and Pic Ningua was not associated with the mean cutout lengths for the trees (Pearson correlation coefficient, $r = -0.0682$, $n = 5$, $P \gg 0.05$, and $r = 0.165$, $n = 10$, $P \gg 0.05$, respectively). The number of stepped-cuts (the square root normalized data) on cutouts was different between Pic Ningua (data for areas A and B combined) and Mt Cindoa when cutout length was accounted for (ANCOVA, F (area) = 9.72, d.f. = 1, $P \ll 0.05$; and F (length) = 3.29, d.f. = 1, $P > 0.05$).

on each occasion. Crows used tools in holes in living and dead wood (Fig. 1a,b) and among the bases of palm and *Freyinetia longispica* (Fig. 1c) leaves, where they captured unidentified prey. I found larvae, spiders, millipedes, weevils, cockroaches, flat-worms, amphipods, isopods, centipedes and earwigs at similar search sites.

Crows searched many sites in different trees with the same tool. The longest period I could stay in visual contact with a tool-using bird was 30 min. I could not establish tool life because birds changed foraging localities frequently, taking their tools with them. Tools I collected from crows I frightened away or those they dropped in my presence were of fresh plant material.

Between foraging episodes, crows often transferred their tools to their feet or placed them in a secure position on their perch, retrieving them with their bills before departing (Orenstein¹⁵ observed a bird put down a tool-like object, then pick it up soon after and fly off with it). Two birds changed trees after putting down their tools, returning within minutes to retrieve them. When feeding on prey, birds left tools in search sites or transferred them to their feet. In the breeding season I observed an adult with both



FIG. 1 Drawings of tool-using behaviour by *C. moneduloides*. Birds are shown searching for prey with hooked-twig tools (held angleways in their bills¹⁵) in the base of *F. longispica* leaves (c) and in the broken-off ends of dead branches (a and b). A stepped-cut tool is shown held longways in a crow's bill (d).

food and a tool in its bill land next to a juvenile, transfer the tool to its feet, feed the juvenile, then pick up the tool and fly off with it.

I collected two types of tool from crows: hooked-twig (Fig. 2c–j), and stepped-cut tools (Fig. 3B, b–j). Hooked-twig tools were manufactured from living secondary twigs. All were stripped of leaves, and usually bark, and had a hook (a section of primary twig) on their wide ends. Stepped-cut tools were cut out along *Pandanus* spp. leaf edges. They had tapered shapes making them pointed but sturdy, and sharp and rigid barbs along their uncut edges which always faced upwards from the narrow ends. Tapered shapes seemed appropriate when birds searched in sites holding these tools longways (Fig. 1d). I only observed birds using the hook end of hooked-twig tools, and the narrow end of stepped-cut ones, in search sites; collected tools I visually inspected were dirty or worn only at the working tips.

Rapid back and forth movements¹⁵ of both tool types were used where prey were under detritus (Fig. 1a). I observed material being hooked out of these sites. Slow and careful movements were used to obtain prey from bases of leaves and holes where it was possible that crows could see their prey (Fig. 1b,c).

I did not observe stepped-cut tool manufacture, but from cutouts I inspected on leaves (Fig. 3) crows cut tools out neatly. Almost-completed tools on leaves (Fig. 3) and associated bill marks indicated birds did this working from the narrow to the wide end. I observed birds manufacture and complete four hooked-twig tools before their use. Birds removed secondary twigs along with part of the primary twig (Fig. 2b). They then flew to perches and transferred them to their feet and spent some minutes working on the hook ends with their bills (I observed hooks on the completed tools). Finally they removed the leaves, and usually bark, from the twig. The numbers of cutouts on tree leaves at Pic Ningua indicated that stepped-cut tools were manufactured there in large numbers (Table 1). Thirty-eight of the 55 tools that I saw being carried by crows and could positively identify were hooked-twig ones.

Tool use by crows was widespread as I collected tools from birds at three other forest sites (Figs 2, 3). These included unhooked-twig tools and an untapered strip of *Pandanus* sp. leaf-edge that birds used in Rivière Bleue Park. The lengths of stepped-cut tool cutouts differed between three areas: two areas at Pic Ningua (Table 1), and one at Mt Cindoa (Table 1). Whereas cutouts had strongly discrete shapes within areas at Pic Ningua, their lengths and the number of steps varied. This indicated that tool standardization¹⁶ was not constrained by raw material or manufacture technique.

That crows used hooks is implied by the manufacture and use of two different tool types that incorporated hooks in a way that should aid prey capture by allowing birds to hook material or prey from search sites. Hook use suggests an appreciation of tool

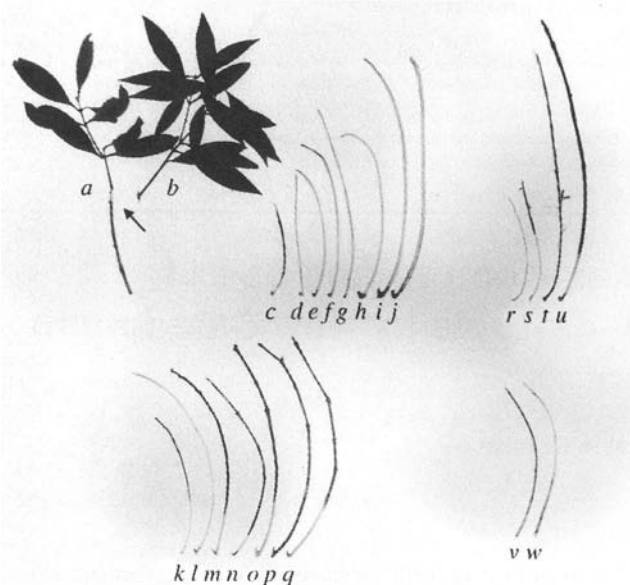


FIG. 2 Twig tools collected from *C. moneduloides* from 1993–95 at Pic Ningua and three other forest sites: Mt Cindoa (6 km from Pic Ningua; 1,100–1,250 m above sea level) and Rivière Bleue Park (60 km southeast of Pic Ningua; 150–200 m) on the mainland, and Maré Island (110 km from the mainland; 0–100 m; crows introduced around 1900). Birds searched for prey with the bottom ends of the tools, which are all hooked except for tools labelled r–t and v–w. Using plant material from *Osmanthus austrocaledonicus* I show how one crow (in tall shrubby vegetation bordering the rainforest) removed twigs for tool manufacture. It removed the secondary twig (b) from the primary one (a; removal position arrowed) with a 'nipping' cut at the intersection of the twigs. This resulted in a 'hook' (the nipped-out section of the primary twig; twigs I broke off did not have hooks) on the secondary twig that needed shaping to match the ones on tools I collected from birds. Pic Ningua tools (c–j) were manufactured from *O. austrocaledonicus* (c), and *Elaeocarpus dognyensis* (d–j). Mt Cindoa tools (k–q) were manufactured from an undescribed *Elaeocarpus* sp. (k–m); an unidentified species (n), and possibly Rubiaceae (o–q). Rivière Bleue Park tools (r–u) were manufactured from *Gymnostoma webbiani* (r–s), *Greslania rivularis* (t, a bamboo flower stem), and Rubiaceae (u). Maré Island tools (v–w) were manufactured from Oleaceae or Verbenaceae. Tool length (measured on a straight line between both ends of a tool) and shaft diameter (taken a few mm from the wide end) were not different between the four sites (Kruskal–Wallis one-way non-parametric ANOVA, d.f. = 20, $F = 0.53$, $P = 0.67$, and one-way ANOVA, $F = 0.83$, $P = 0.50$, respectively). For scale, tool i is 19.2 cm in length.

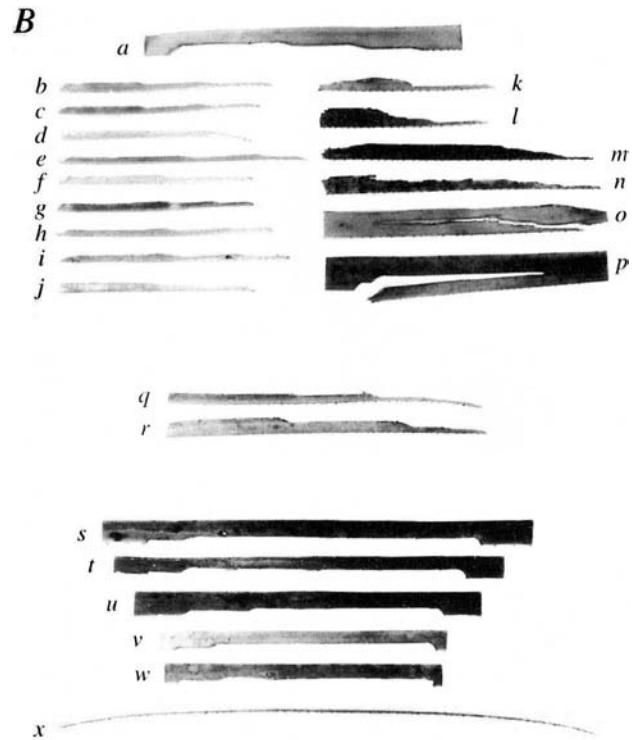


FIG. 3A, An almost-completed stepped-cut tool on a *Pandanus* sp. leaf (leaf width at the cutout is ~ 7.5 cm) from the leaf underside looking skywards (leaf tips are at the top of the photograph). All cutouts I observed had their narrow ends oriented towards the trunk. B, Stepped-cut tools collected from *C. moneduloides* and leaf material collected from *Pandanus* spp. trees at Pic Ningua and two other forest sites (Mt Cindoa, and Rivière Bleue Park) from 1993–95. Leaf-edge barbs (bottom sides) on material a–r face left, and rightwards on s–x (the narrow ends on some tools bowed when drying). Pic Ningua material: an example of a stepped-cut tool cutout (a); stepped-cut tools (all 3-stepped) collected from crows (b–j; mean length, 14.90 cm; s.e.m. = 0.410); 3 completed tools (k–m) I found under one tree (respective cutouts were on its leaves), and two almost-completed ones on its leaves (n, o; tool n was removed from the leaf as it was just attached), which were noticeably misshapen compared to tools b–j; an example (p) of leaf-edge strips which were commonly cut (bill marks were often present) and ripped (always towards the leaf bases) on leaves at Pic Ningua and Mt Cindoa, but less so at the Rivière Bleue Park. Mt Cindoa

material: the tool I collected from a crow (q; length, 21.4 cm) was similar in shape to the one I found among leaves in a tree there (r; length, 22.0 cm). Rivière Bleue Park material: five cutouts (all 2-stepped) I collected off three trees (s–w; mean length of missing leaf edge = 19.62 cm, s.e.m. = 1.106); an unstepped-cut tool I collected from a crow (x; length = 37.8 cm; working tip at left).

functionality¹⁷, and a tool kit suggests different tools for different tasks¹⁸.

Crow tool manufacture had three features new to tool use in free-living nonhumans, and that only first appeared in early human tool-using cultures after the Lower Palaeolithic^{6,7}: a high degree of standardization, distinctly discrete tool types with definite imposition of form in tool shaping, and the use of hooks. □

Development of identical orientation maps for two eyes without common visual experience

Imke Gödecke & Tobias Bonhoeffer*

Max-Planck Institute for Psychiatry, Am Klopferspitz 18A, 82152 München-Martinsried, Germany

In the mammalian visual cortex, many neurons are driven binocularly and response properties such as orientation preference or spatial frequency tuning are virtually identical for the two eyes¹. A precise match of orientation is essential in order to detect disparity and is therefore a prerequisite for stereoscopic vision. It is not clear whether this match is accomplished by activity-dependent mechanisms together with the common visual experience normally received by the eyes^{2,3}, or whether the visual system relies on other, perhaps even innate, cues to achieve this task^{4–7}. Here we test whether visual experience is responsible for the match in a reverse-suturing experiment in which kittens were raised so that both eyes were never able to see at the same time. A comparison of the layout of the two maps formed under these conditions showed them to be virtually identical. Considering that the two eyes never had common visual experience, this indicates that correlated visual input is not required for the alignment of orientation preference maps.

* To whom correspondence should be addressed.

Received 13 June; accepted 15 November 1995

1. Chevalier-Skolnikoff, S. *Behav. Brain Sci.* **12**, 561–627 (1989).
2. Parker, S. & Gibson, K. J. *hum. Evol.* **6**, 623–641 (1977).
3. Beck, B. B. *Animal Tool Behaviour: The Use and Manufacture of Tools by Animals* (Garland, New York, 1980).
4. Morse, D. H. *The Wilson Bulletin* **80**, 220–223 (1968).
5. Millikan, G. C. & Bowman, R. I. *Living Bird* **6**, 23–41 (1967).
6. Mellars, P. in *The Human Revolution* (eds Mellars, P. & Stringer, C.) 338–365 (Princeton University Press, New Jersey, 1989).
7. Yellen, J. E., Brooks, A. S., Cornelissen, E., Mehlman, M. J. & Stewart, K. *Science* **268**, 553–556 (1995).
8. Delacour, J. *Guide des oiseaux de la Nouvelle-Calédonie et ses Dépendances* (Delachaux and Niestlé, Neuchâtel, 1966).
9. Reid, J. B. *Anim. Behav.* **30**, 1212–1216 (1982).
10. Jones, T. B. & Kamil, A. C. *Science* **180**, 1076–1078 (1973).
11. Grobecker, D. B. & Pietsch, T. W. *Auk* **95**, 760–761 (1978).
12. Zach, R. *Behaviour* **68**, 106–117 (1979).
13. James, P. C. & Verbeek, N. A. M. *Ardea* **72**, 207–215 (1984).
14. Gayou, D. C. *The Wilson Bulletin* **94**, 593–594 (1982).
15. Orenstein, R. I. *Auk* **89**, 674–676 (1972).
16. Napier, J. in *Classification and Human Evolution* (ed. Washburn, S. L.) 3rd edn 178–189 (Aldine, Chicago, 1969).
17. Caron, A. J., Caron, R. F. & Antell, S. E. *Dev. Psychol.* **24**, 620–627 (1988).
18. McGrew, W. C. *Chimpanzee Material Culture: Implications for Human Evolution* (Cambridge University Press, Cambridge, 1992).

ACKNOWLEDGEMENTS. I thank C. Veltman for her advice, ideas and bibliographic help; E. Grant for Figure 1; M. Ménézo (ORSTOM) for use of computer facilities; and J.-M. Veillon for identification of plant material. Work at Rivière Bleue Park and Pic Ningua was courtesy of the Parks and Reserves Department. This study was undertaken during my kagu research funded by BirdLife International and The Royal Society for the Protection of Birds. My Maré Island visit was self-funded.

We used the reverse-suturing protocol as a suitable method for artificially separating the formation of orientation maps for the two eyes. This enabled us to determine the extent to which visual experience establishes the correspondence between these maps: kittens were raised with exclusively monocular experience up to the fifth postnatal week. We then used optical imaging based on intrinsic signals to record the orientation preference map in area 18 of the visual cortex by stimulating the open eye. Subsequently, the eyes were reverse-sutured and the newly opened eye was allowed to establish an orientation preference map. Seven to fourteen days after reverse-suturing, we imaged the same cortical region, stimulating the now open eye and compared the new map with the previously recorded one.

Figure 1 shows activity maps obtained from one kitten stimulated with moving gratings of four different orientations. A comparison of the activity maps obtained by stimulation of the left eye at postnatal day (PND) 35 (left column) and the activity maps 13 days after reverse-suturing (right column, PND 48, right eye) clearly shows that the iso-orientation maps formed by the initially deprived eye are very similar to those formed by the eye that was visually experienced from birth. Some minor differences can be observed in the precise structure of the maps. Bearing in mind, however, that the first activity maps were obtained through an intact dura to minimize damage to the cortex, the correspondence is quite remarkable.

Although it is well established that, after this period of deprivation, tuned single-cell responses and consequently orientation maps are virtually absent for the deprived eye⁸⁻¹⁰, we verified that stimulation of the initially closed eye did not result in an orientation map. In one of the experiments, therefore, we also determined the orientation preference map of the initially closed eye immediately after reverse suturing (Fig. 2). Although stimulation of the visually experienced eye leads to patterned activity in

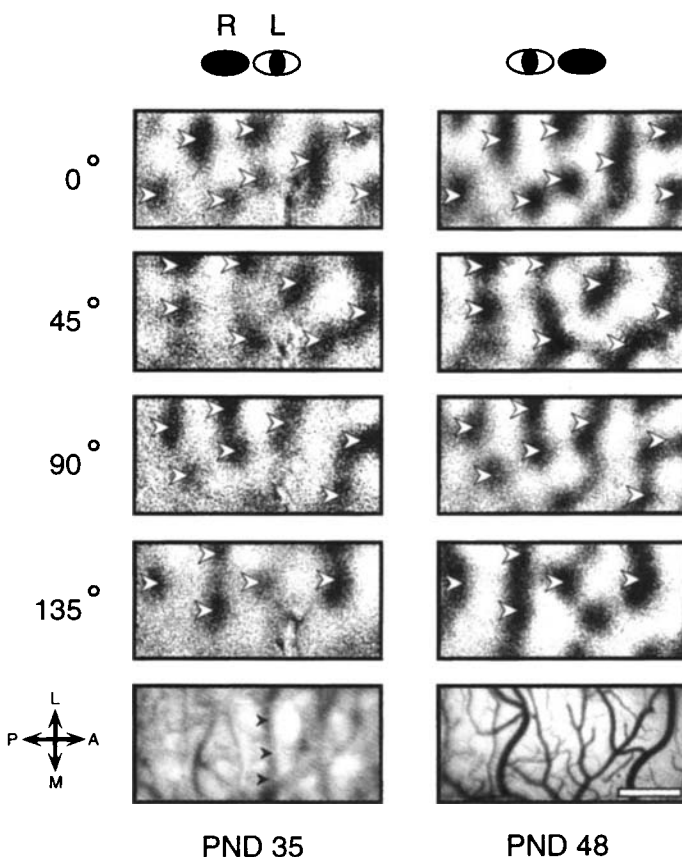
the cortex (left column), absolutely no sign of activity maps could be detected for the deprived eye during the initial recording at PND 34 (middle column). Nevertheless, after eight days of visual experience, the same eye was able to drive cortical neurons and the orientation maps (right column) were again virtually identical to the maps of the initially open eye.

The complete layout of orientation preference is best visualized by colour-coded angle maps, in which different colours represent the preferred stimulus orientation¹¹⁻¹⁴. Figure 3 shows such maps calculated from the data of a third experiment of the kind already described. In this case too, there is a striking similarity between the orientation preference maps for the two eyes that had never received common visual stimulation. Minor differences between the maps, such as a slight shift of some pinwheel centres, are probably due to technical limitations such as a slight distortion of the cortical surface by a change in intracranial pressure, rather than reflecting real differences in the layout of these maps. In fact, the root-mean-square (r.m.s.) difference between the two maps that developed independently (27.6°) comes close to the value for technical reproducibility calculated from the data obtained in one recording session (22.9°; see also ref. 14).

In order to quantify further the similarity of the orientation maps for the left and right eye, we calculated two-dimensional cross-correlations. This enabled us to obtain a useful correlation coefficient, even if the two maps were not perfectly aligned (the maximum was in this case slightly shifted from the origin). Figure 4 uses bar-plots to display the correlation coefficients for orientation maps obtained for left and right eye from three kittens, in which the data were of superior quality so that this type of quantitative analysis could be performed. Correlation coefficients for maps obtained from the two eyes of the same animal are markedly larger than the correlation coefficients resulting from the comparison of orientation maps from left and right eye of

FIG. 1 Iso-orientation preference maps from the left and right eye of a kitten raised with alternate eye occlusion. Panels on the left side show the iso-orientation maps obtained from the left eye at PND 35. After the recording session, the eyes were reverse-sutured and the newly opened eye was allowed independently to gain visual experience until PND 48. At this point, the newly formed iso-orientation maps were recorded by stimulating the right eye (right column). To facilitate comparison of the orientation maps from left and right eye, arrowheads are placed in corresponding positions in the right and left activity maps. This reveals the striking similarity between the orientation maps formed independently by both eyes. The blood-vessel patterns of the different recording sessions are displayed at the bottom of the columns. Taking into account that the left images were obtained through the dura, the correspondence of the blood vessels is quite good apart from one large vessel running from top to bottom in the middle of the left picture (dark arrowheads). This is a dural vessel and can therefore not be seen in the second experiment (lower right panel), in which the dura was removed. Scale bar, 1 mm.

METHODS. In all animals, the right eye was sutured before the time of natural opening. The kittens were anaesthetized with a combination of ketamine (Ketanest, Bayer; 20–40 mg kg⁻¹) and xylazine (Rompun, Bayer; 2–4 mg kg⁻¹). The right eyelid was pried open, its margin trimmed and sutured shut. The seam was carefully observed every day in order to close developing holes before the kitten was able to see through them. Chamber implantation: the kitten was initially anaesthetized with ketamine/xylazine, intubated, and artificially respiration (30–50% O₂, 50–70% N₂O, 0.7–1.2% halothane) during surgery. ECG, arterial oxygen saturation (SpO₂), expired CO₂, and rectal temperature were monitored throughout the experiment. The skull was opened above area 18 around Horsley-Clark A2. A titanium chamber was implanted, filled with silicon oil and closed with a glass coverslip. To prevent infection the dura was always covered with a thin layer of agar supplemented with antibiotics (chloramphenicol (Paraxin, Bayer); 10 mg). After implantation, kittens were allowed to recover from anaesthesia before being returned to their mother and littermates. Details of the optical imaging technology can be found elsewhere^{13,14}. Briefly, the cortical surface was illuminated with light of 605 nm wavelength. An anaesthetized animal (N₂O/O₂/halothane anaesthesia) was stimulated with moving square wave gratings of different orientations (2 cycles per second, 0.15 cyc per deg), projected onto a frosted glass screen 1 m in front of the properly refracted eyes of the animal. Images were acquired from the exposed part of the cortex with a slow-scan CCD-camera. Activity maps were



calculated by dividing the images obtained for one orientation by the sum of the images obtained for all orientations (cocktail-blank¹⁴, the homogeneous activation of the cortex).

different kittens. The latter demonstrates that a general structural similarity between orientation maps can not account for the high correlation found between maps obtained from the two eyes of one kitten. The quantitative analysis of our data then supports the conclusion drawn from the qualitative comparison of the orientation maps for left and right eye, namely that there is a striking similarity between the maps, even if the eyes had never received correlated visual experience.

In two further experiments, we only obtained data of substantially lower quality in the second imaging session. Consequently, we could not use these data to perform a quantitative analysis as presented in Fig. 4. These results were analysed qualitatively, using a blind evaluation procedure. In these cases too, we found clear evidence for a high correlation between the maps from the two eyes.

Earlier studies reported a strong influence of visual experience on the shaping of response properties of cells in the visual cortex^{2,3,9}. Our data indicate that the influence of visual experience on the layout of cortical maps (and therefore presumably also on the response properties of single cells) is weaker than previously thought¹⁵. This is, of course, not to imply that visual experience is not important for maintenance and development of cortical maps. Evidently, deprivation of vision in one eye causes the responses, and consequently also the map for this eye, to vanish (Fig. 2 and refs 8, 10). The important point here is that the layout of this map is not altered by this kind of experimental manipulation (see also ref. 10). The fact that different visual experience cannot cause a different orientation map to form for the second eye does, of course, not rule out that the layout of the first orientation map might have been influenced by visual experience. It has to be borne in mind that the second eye encounters a cortex in which an

orientation map had already been present—albeit for the other eye and at an earlier time. It has, for instance, been proposed that orientation preference of cortical neurons is largely determined by intracortical connections^{16–18}. Specifically, it has recently been shown that intracortical connections can be anisotropic along the axis of preferred orientation¹⁹. If such anisotropies are present for the first eye and if both eyes' retinotopic maps are already in register, the orientation preference set up by the first eye could be read out by the second eye from these anisotropic connections²⁰. But there is also an alternative explanation for our finding. It has been suggested that orientation preference might be innately determined^{4,5}, a conclusion based mainly on the fact that newborn animals already have orientationally tuned cells. This reasoning was undermined, however, by experiments demonstrating travelling activity waves in the embryonic retina²¹. These wavefronts could mimic contours in the visual scenery and therefore, together with an activity-dependent mechanism, they could explain orientation maps in newborn animals—such maps would not be innate in the strict sense. The results presented here cannot be explained by retinal waves, which would have to be synchronized between the two eyes. Therefore our data can also be interpreted as supporting the idea that there is indeed a truly innate component to the layout of orientation maps. In fact, if one assumes that orientation selectivity of single neurons is chiefly determined by the geometry of geniculocortical afferents^{22–25}, rather than by intracortical circuitry, our results must indicate that orientation preference is not an acquired, but a predetermined response property.

Further experiments will be necessary to distinguish between these two possibilities and to find out how the match of orientation preference is accomplished. What we have shown, however, is that common visual experience is not required to achieve this task: it

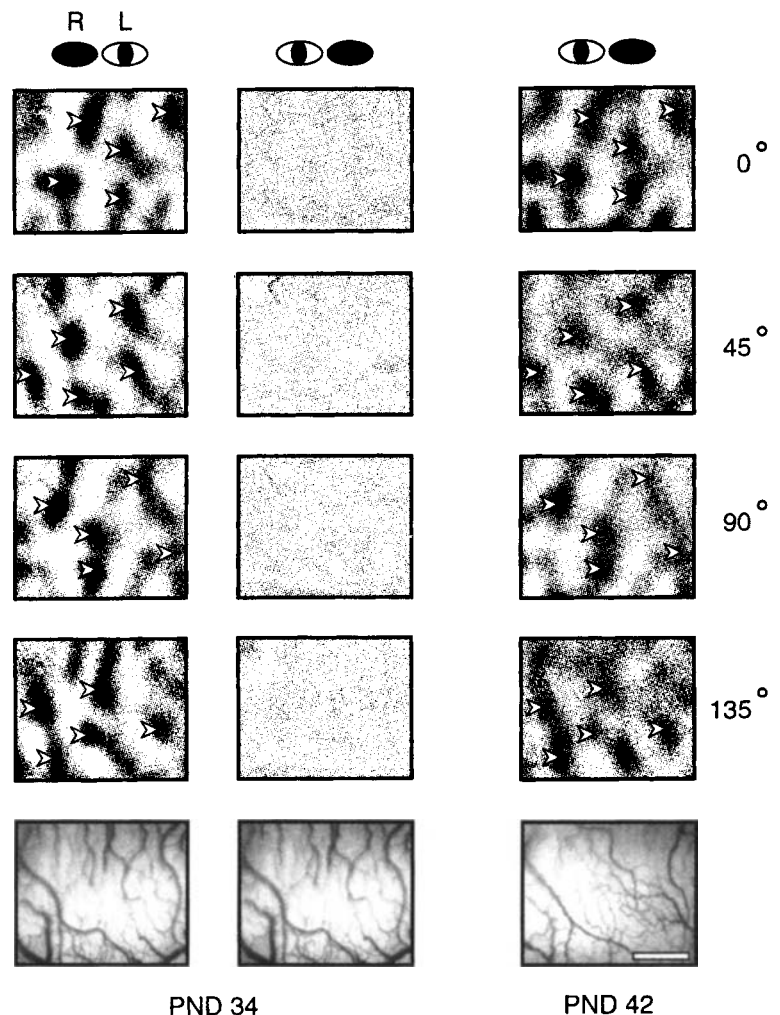


FIG. 2 Control experiment showing that the closed eye does not form an orientation map. The left column shows four activity maps obtained by stimulating the (visually experienced) left eye at PND 34. Within the same recording session, the eyes were reverse-sutured and activity maps were recorded for the newly opened eye. These maps, shown in the middle column, are essentially flat. Eight days later (at PND 42) the same eye had established clear orientation maps, which again were virtually identical to the initial maps. To facilitate comparison, arrowheads are placed in corresponding places of the maps. Scale bar, 1 mm.

FIG. 3 Colour-coded orientation preference maps for a third kitten raised with alternate eye occlusion. For this animal, activity maps were recorded at PND 34 (left eye) and PND 47 (right eye). In both cases, responses to all stimulus orientations were summed vectorially. The preferred orientations for every cortical site are displayed using a discrete colour code, shown at the right of the figure. Areas responding best to horizontal bars are coded in yellow, areas responding best to vertical bars in blue, and so on. Arrowheads were positioned to mark some of the pinwheel¹³ centres in the left orientation preference map. The same pattern of arrowheads was superimposed onto the right map to facilitate comparison between the two maps. Again, the angle maps from the left and right eye show a remarkable similarity. Blood-vessel patterns of the corresponding recording sites are shown at the bottom of the figure. The image for PND 34 is less crisp as recording and blood-vessel patterns were obtained through the intact dura. Scale bar, 1 mm.

METHODS. Angle maps were computed by summing responses to all stimulus orientations vectorially, taking the magnitude of the responses as vector lengths and the stimulus orientation as their vector angle¹¹⁻¹⁴. The angle of the resulting vector was then displayed using the discrete pseudo-colour code shown.

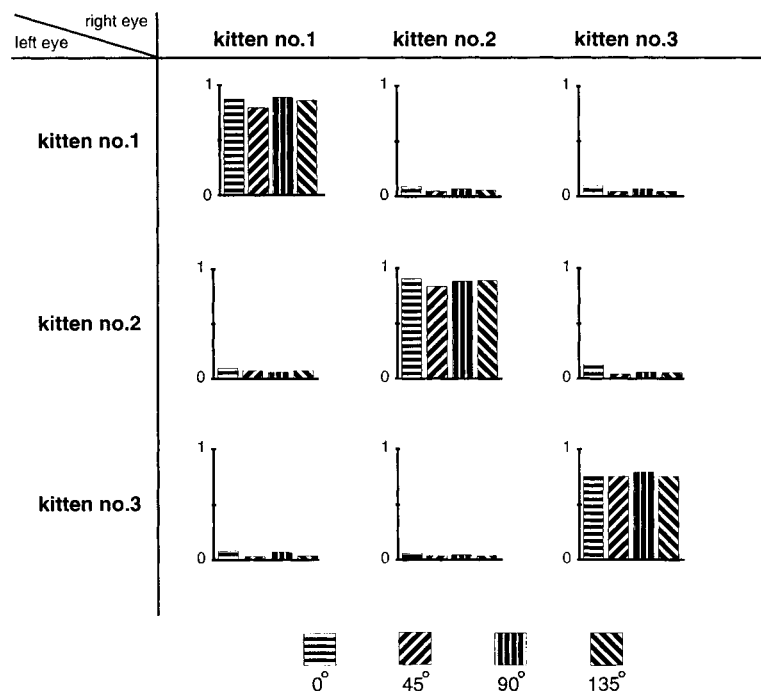
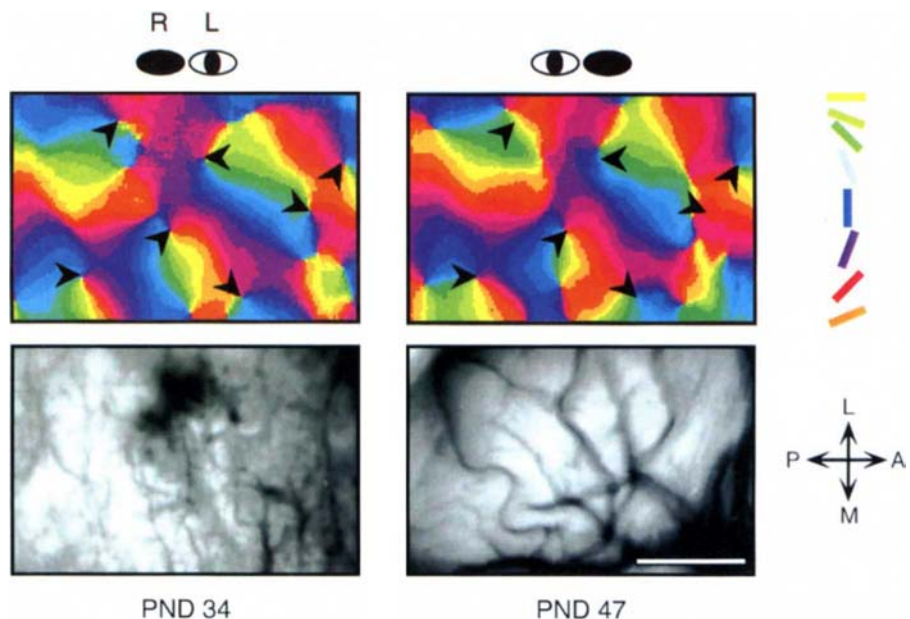


FIG. 4 Cross-correlation coefficients for maps from different animals. Two-dimensional cross-correlations were calculated for maps obtained with the same stimulus orientation for the left and right eyes in three different animals. Each cell contains a bar plot in which the lengths of the bars display the maximum correlation coefficients for the iso-orientation maps of the respective stimulus orientations (the orientation of the stimulus is given by the orientation of the hatching). There is a high degree of similarity between maps obtained from the different eyes of one animal, whereas the similarity of maps from different animals is low. The correlation coefficients shown for the different animals (the off-diagonal values) are an upper estimate for the similarity, because the computer searched for the maxima of the 2-dimensional cross-correlation. In all off-diagonal cases, the corresponding minima were of similar magnitude, indicating that the actual cross-correlation was close to zero.

METHODS. Activity maps were aligned with the help of the blood-vessel patterns obtained before the respective recording sessions. As some of the recordings were performed through the intact dura, a perfect alignment was sometimes difficult to achieve. We therefore calculated two-dimensional cross-correlations in which the correlation coefficient between the maps was calculated for every offset between $-800\ \mu\text{m}$ and $800\ \mu\text{m}$ (-40 to $+40$ pixels). The maximum of the correlation is shown here.

seems to be so important to the visual system that the orientation maps for the two eyes match, that it does not rely just on visual experience to achieve this. □

Received 13 September; accepted 16 November 1995.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol.* **160**, 106–154 (1962).
- Blakemore, C. & Van Sluyters, R. C. *J. Physiol.* **248**, 663–716 (1975).
- Barlow, H. B. & Pettigrew, J. D. *J. Physiol.* **218**, 98–100 (1971).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **26**, 994–1002 (1963).
- Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **158**, 307–318 (1974).
- Albus, K. & Wolf, W. *J. Neurophysiol.* **348**, 153–185 (1984).
- Braastad, B. O. & Heggelund, P. J. *J. Neurophysiol.* **53**, 1158–1178 (1985).
- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003–1017 (1963).
- Imbert, M. & Buisseret, P. *Exp Brain Res.* **22**, 25–36 (1975).
- Kim, D.-S. & Bonhoeffer, T. *Nature* **370**, 370–372 (1994).
- Blasdel, G. G. & Salama, G. *Nature* **321**, 579–585 (1986).

- Ts'o, D. Y., Frostig, R. D., Lieke, E. E. & Grinvald, A. *Science* **249**, 417–420 (1990).
- Bonhoeffer, T. & Grinvald, A. *Nature* **353**, 429–431 (1991).
- Bonhoeffer, T. & Grinvald, A. *J. Neurosci.* **13**, 4157–4180 (1993).
- Movshon, J. A. *J. Physiol.* **261**, 125–174 (1976).
- Sillito, A. M., Kemp, J. A., Berardi, J. & Berardi, N. *Brain Res.* **194**, 517–520 (1980).
- Crook, J. M. & Eysel, U. T. *J. Neurosci.* **12**, 1816–1825 (1992).
- Schwarz, C. & Bolz, J. *J. Neurosci.* **11**, 2995–3007 (1991).
- Fitzpatrick, D., Zhang, Y., Schofield, B. R. & Muly, E. C. *Soc. Neurosci. Abstr.* **19**, 179.2 (1993).
- Bonhoeffer, T. & Kim, D.-S. in *Challenges and Perspectives in Neuroscience* (ed. Ottoson, D.) 117–138 (Elsevier, Oxford 1995).
- Meister, M., Wong, R. O. L., Baylor, D. A. & Shatz, C. J. *Science* **252**, 939–943 (1991).
- Tanaka, K. *J. Neurophysiol.* **49**, 1303–1318 (1983).
- Chapman, B., Zahs, K. R. & Stryker, M. P. *J. Neurosci.* **11**, 1347–1358 (1991).
- Ferster, D. *J. Neurosci.* **6**, 1284–1301 (1986).
- Reid, R. C. & Alonso, J. M. *Nature* **378**, 281–284 (1995).

ACKNOWLEDGEMENTS. We thank F. Brinkmann for technical assistance and D. Fitzpatrick for comments on an earlier version of the manuscript. The work was partly supported by a grant from the Human Frontier Science Program Organisation.

Conservation of hippocampal memory function in rats and humans

M. Bunsey* & H. Eichenbaum†

Center for Behavioral Neuroscience, State University of New York at Stony Brook, New York 11794-2575, USA

THE hippocampus is critical to declarative memory in humans¹. This kind of memory involves associations among items or events that can be accessed flexibly to guide memory expression in various and even new situations²⁻⁴. In animals, there has been controversy about whether the hippocampus is specialized for spatial memory^{5,6} or whether it mediates a general memory function^{3,4}, as it does in humans. To address this issue we trained normal rats and rats with hippocampal damage on non-spatial stimulus-stimulus associations, then probed the nature of their memory representations. We report here that normal rats demonstrated two forms of flexible memory expression, transitivity, the ability to judge inferentially across stimulus pairs that share a common element, and symmetry, the ability to associate paired elements presented in the reverse of training order. Rats with neurotoxic damage limited to the hippocampus demonstrated neither form of flexible expression, indicating that non-spatial declarative processing depends specifically on the hippocampus in animals as it does in humans.

Declarative memory in humans is commonly assessed by the verbal paired-associate task in which subjects study arbitrarily paired words and are subsequently tested by presenting the first word of each pair and assessing recall of the second. Exploiting rodents' natural foraging strategies that use olfactory cues, we developed an odour-guided version of the paired-associate task to assess declarative memory in rats. Animals were trained with stimuli that consisted of distinctive odours added to a mixture of ground rat chow and sand through which they dug to obtain buried cereal rewards (Fig. 1). On each paired-associate trial one of two sample odours initially presented was followed by two choice odours, each assigned as the 'associate' of one of the samples and baited only when preceded by that sample. Thus, much as in human paired-associate learning, the rats were required to form specific associations between stimuli and to identify the associated choice stimulus that followed each sample. Following training on two sets of overlapping odour-odour associations, subsequent probe tests were used to characterize the extent to which learned representations supported transitivity and symmetry.

FIG. 2 Brain sections from a representative animal with selective damage to the hippocampus (Ammon's horn and dentate gyrus). The rostral (top) section shows complete loss of dorsal Ammon's horn and only a remnant of the dentate gyrus. The mid-coronal (middle) section shows nearly complete cell loss in all fields of the dorsal and ventral hippocampus. The caudal (bottom) section shows a partial lesion restricted to the posterior dentate gyrus and hilus. The intact fimbria-fornix can be seen in the rostral two sections, and the intact subiculum can be seen in the caudal section. Mid-coronal and caudal sections show the spared entorhinal and perirhinal cortical areas. Pipette penetrations and local thinning in parietal cortex surrounding the injection tract can be seen above the hippocampal lesion in the rostral two sections. Minor cortical damage and thinning was observed above the pipette track in all hippocampal animals, as well as in controls, and was not related to performance. There was no observable cell loss in entorhinal cortical layers II/III that project to the hippocampus. A similar pattern and amount of hippocampal damage, as well as sparing of surrounding areas, was observed in all subjects of the hippocampal group. METHODS. Hippocampal cell destruction in male Long-Evans rats was produced by 14 microinjections (0.04–0.08 μ l) of the neurotoxin ibotenic acid into each hemisphere according to the methods and coordinates of ref. 31 (see ref. 32 for more detailed analyses of similar lesions). Subjects

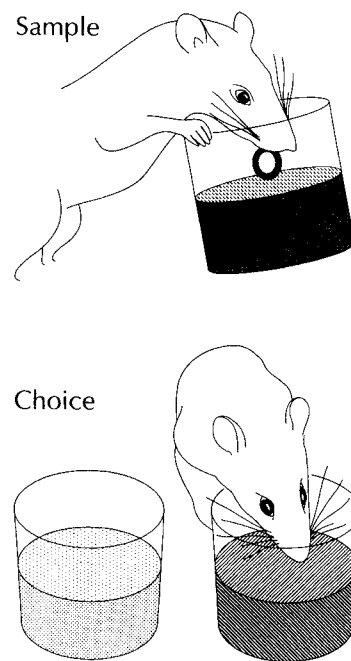
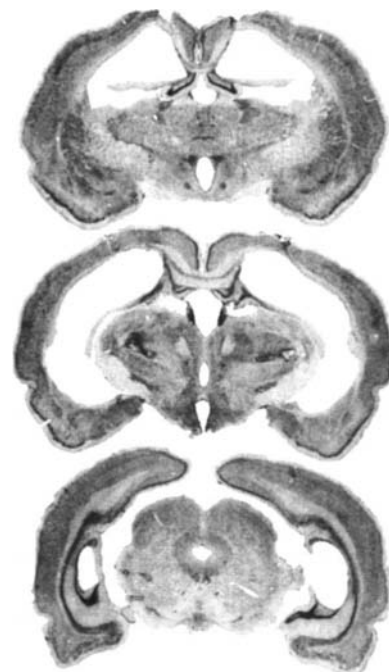


FIG. 1 Training on odour-odour paired associates. Each training trial consisted of two phases. In the sample phase, the subject was presented with a cup containing a scented mixture of sand and ground rat chow with a buried reward. In the subsequent choice phase, two scented choices were presented. Both choice items involved odours that were different from the sample odour, and which item was baited depended on the identity of the sample.

METHODS. All training and testing was performed in the home cage. Rats were first 'shaped' to obtain bits of buried sweet cereal by digging through a 50:50 clean sand and ground rat chow mixture contained in a 6.5 cm diameter $\times$ 6 cm tall clear plastic cup. In subsequent training the sample cup always had a reward buried at the bottom of the cup. In the choice phase, the first cup in which a subject began to dig was recorded as the response selection, though subjects were allowed to dig in both cups until they retrieved the reward.



were allowed to recover for 3.5 weeks before training. Following testing, brains were fixed in formalin solution, frozen, sliced at 30–40 μ m in the coronal plane, and stained with thionin.

Twenty rats received neurotoxic cell destruction selectively in the hippocampus and dentate gyrus (Fig. 2) or sham control operations. After recovery, training and testing proceeded in four distinct phases (Fig. 3a). First, subjects were trained on a set of two paired associates, followed by a second set in which the

previous choice odours served as samples, and two new odours were assigned as choice items. Then, to test transitivity, subjects were continued on the paired associates from both training sets, to which were added occasional probe trials comprising a sample odour from set 1 and unbaited choice odours from set 2. Time

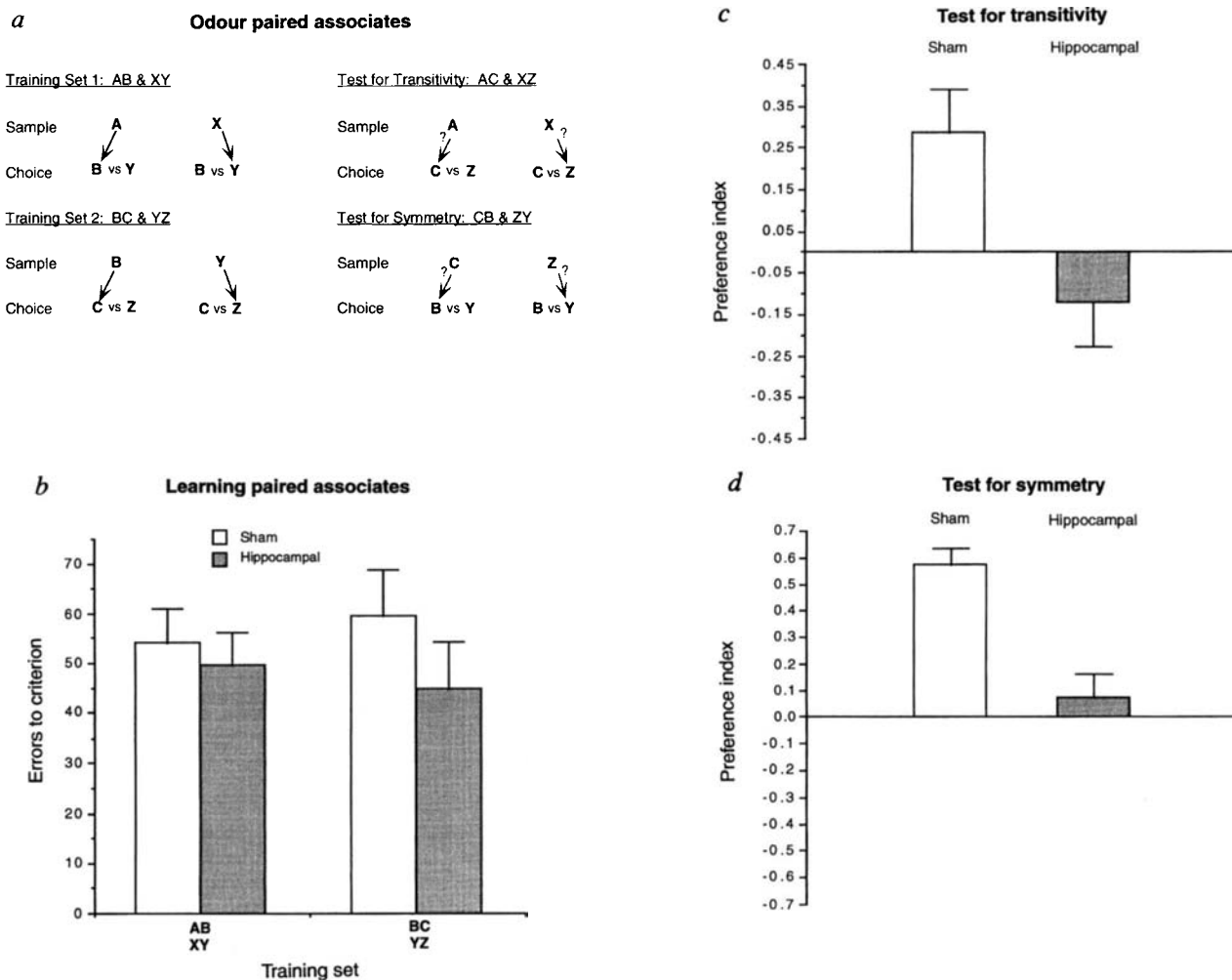


FIG. 3a, Schematic diagram of paired-associate training and probe testing. Letters represent odour stimulus items; arrows without question marks indicate trained pairings, whereas arrows with question marks indicate expected transitive and symmetrical choices. b, Errors to criterion (+ s.e.) on acquisition of the two sets of paired associates for sham-operated ($n = 12$) and hippocampal ($n = 8$) subjects. ANOVA revealed no significant differences in performance between groups ($F(1, 18) = 1.20$) or between the two sets of paired associates ($F(1, 18) = 0.03$), and no significant group by training set interaction ($F(1, 18) = 0.38$). c, Preference indices (+ s.e.) on the test for transitive inference. On each probe trial a preference score was calculated as $(X - Y)/(X + Y)$, where X and Y were the digging times in the transitive and alternate choices, respectively. The preference index for each subject was the average of these scores across all probe trials. Sham-operated controls preferred the transitive choice more than predicted by chance ($F(1, 11) = 2.67$, $P = 0.02$) and their preference differed significantly from that of the hippocampal group ($F(1, 18) = 6.398$, $P = 0.02$). The hippocampal group showed no significant preference ($F(1, 7) = -1.06$). Average total digging times on probe trials did not differ between groups (control = 3.97 s, hippocampal = 3.54 s; $t = 0.74$). d, Preference indices (+ s.e.) on the test for symmetrical expression, calculated as in the transitivity test but using one score for each subject based on the first response on each reinforced trial. The sham-operated subjects ($n = 6$) strongly preferred the symmetrical choice ($F(1, 5) = 8.95$, $P = 0.0003$) and their preference differed significantly from the hippocampal group ($F(1, 8) = 21.36$, $P = 0.0001$). The hippocampal group ($n = 4$) showed no significant preference ($F(1, 3) = 0.775$). METHODS. Training on paired-associate sets. The sample odours for the

first set of paired associates were cocoa (7% by weight) or 5% turmeric and the choice odours were 15% coffee and 30% salt. The odour pairing assignments were counterbalanced across subjects, and the left-right positions of the choice cups followed a pseudorandom trial sequence. Each day 14 trials were conducted until subjects reached a criterion of 11 correct on 3 out of 4 consecutive days. In preparation for non-rewarded probe trials, on 2–3 trials per day neither choice cup was baited; instead the reward was dropped on the surface of the sand mixture following the choice response. Performance on these trials confirmed that subjects were not solving the task by directly detecting buried rewards on trials when the cups were baited. The sample odours on the second set of paired associates were coffee and salt, the same odours as the choice items from the first set. For choice items on this set, coffee was paired with 4% onion, and salt was paired with 2% nutmeg. Transitive inference test. On one test day, subjects were presented with 10 trials from both sets of paired associates plus 4 intermixed probe trials, each composed of one of the samples from the first set and both choice items from the second set. The choices on these trials were unbaited and the experimenter recorded the amount of time spent digging in each of the choice cups for a combined maximum of 10 s or until digging ceased for 10 s. Symmetry test. Subjects were retrained to criterion on set 1 using the same protocol but the sample and choice items were reversed in their order of presentation, that is, the former sample items were now the choices and vice versa. Then, using the same protocol, they were given one 14-trial reinforced session using set 2, but also with the sample and choice items reversed in order of presentation. On each trial the first digging response was recorded and these choices were used to calculate the preference index.

spent digging at each choice odour was measured and transitivity was indexed by preferential digging in the choice cup associated only indirectly with the sample. Finally, half of the subjects were given a symmetry test. In constructing this test we took into consideration previous evidence that consistent symmetry is not observed immediately in animals^{7,8} and our own finding of a general slowing of performance during repeated non-reinforced transitivity trials. To address these issues, we first exposed subjects to the symmetry protocol by retraining them on set 1 with the sample and choice items reversed. Then we presented a single symmetry test session on set 2 using reinforced trials.

Control rats learned both sets of paired associates rapidly and hippocampal damage did not affect acquisition rate, consistent with recent reports on stimulus-stimulus association learning in rats and monkeys⁹⁻¹¹ (Fig. 3b). Controls showed strong transitivity across the sets with a preference index corresponding roughly to 2:1 digging in favour of choice items indirectly associated with the presented sample (Fig. 3c). By contrast, rats with hippocampal damage were severely impaired in that they showed no evidence of transitivity. In subsequent retraining on set 1 with reversed sample and choice items, as expected we did not observe symmetry in either group immediately, but nearly all rats showed savings and the amount of savings did not differ significantly between groups (data not shown). In the symmetry test on set 2, controls had a preference index corresponding roughly to 3:1 choices in the direction of the symmetrical association (Fig. 3d). By contrast, rats with hippocampal damage again were severely impaired, showing no significant capacity for symmetry.

The present results support the notion that animals organize non-spatial memory representations so that they can be assessed flexibly and used to guide inferences from memory^{12,13}. These capabilities offer strong links between characterizations of spatial memory in rats and declarative memory in both animals and humans¹⁴. Tolman's¹⁵ view of cognitive mapping highlighted rats' abilities to use inferential behaviour in solving maze problems. Our observation that rats formed integrated representations of overlapping stimulus associations, and spontaneously accessed this organization when challenged with new pairings or training order, is entirely consistent with Dickinson's¹⁶ qualifications of declarative memory from animal learning theory. The present findings also provide an extension to animals of classic views on human memory, such as the description by James¹⁷ of description of 'memory' as involving an elaborated network of associations that can be applied across a broad range of situations, as distinct from 'habits' that depend on rigid associative sequences. Our findings are also entirely consistent with present-day characterizations of human declarative memory, such as Cohen's¹⁴ description of declarative memory as 'promiscuous' in its accessibility by new routes of expression.

In addition, the present findings also support the view that the hippocampus is a critical part of a brain system for mediating non-spatial as well as spatial declarative memory in animals. In particular these data speak to the current issue of which components of the hippocampal region are critical to this kind of memory¹⁸⁻²³, showing that damage limited to the hippocampus itself can produce a severe deficit in non-spatial declarative memory processing in animals, as it does in humans²⁴.

Notably human amnesics typically show severe deficits on standard verbal paired-associate learning, whereas in the present study rats with hippocampal damage successfully acquired odour associations. However, human amnesics can succeed in learning stimulus-stimulus associations when training procedures emphasize implicit memory expression^{25,26,33}. Our previous studies on animals also indicate that spatial²⁷, as well as non-spatial²⁸, learning succeeds without hippocampal function, but only under restricted conditions that minimize demands for flexible representation. Previous studies have emphasized the inflexibility of learning by amnesic humans²⁵ and animals^{29,30}, but these accounts have focused on amnesics' rigidity against changing a prepotent behavioural response. By contrast, the present experiment focuses

on flexibility of access to memories as the central feature of declarative memory representation^{31,34}. Thus, in both humans and animals, stimuli can be associated independently of hippocampal function but the establishment of representations that can be expressed indirectly and inferentially is critically supported by the hippocampus. By this view the most prominent examples of declarative expression, conscious recollection in humans and spatial mapping in rodents, are both manifestations of the same fundamental cognitive function mediated by the hippocampus across species. □

Received 18 September; accepted 2 November 1995.

1. Cohen, N. J. & Squire, L. R. *Science* **210**, 207-210 (1980).
2. Cohen, N. J. in *The Neuropsychology of Memory* (eds Butters, N. & Squire, L. R.) 83-103 (Guilford, New York, 1984).
3. Eichenbaum, H., Otto, T. & Cohen, N. J. *Behav. neural Biol.* **57**, 2-36 (1992).
4. Squire, L. R. (1992) *Psychol. Rev.* **99**, 195-231 (1992).
5. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford Univ. Press, 1978).
6. Morris, R. G. M., Garrud, P., Rawlins, J. N. P. & O'Keefe, J. *Nature* **297**, 681-683 (1982).
7. Sidman, M. et al. *J. exp. Anal. Behav.* **37**, 23-44 (1982).
8. D'Amato, M. R., Salmon, D. P., Loukas, E. & Tomie, A. *J. exp. Anal. Behav.* **44**, 35-47 (1985).
9. Murray, E. A., Gaffan, D. & Mishkin, M. *J. Neurosci.* **13**, 4549-4561 (1993).
10. Davidson, T. L., McKernan, M. G. & Jarrard, L. E. *Behav. Neurosci.* **107**, 227-234 (1993).
11. Cho, Y. & Kesner, R. P. *Behav. Brain Res.* **67**, 91-98 (1995).
12. McGonigle, B. & Chalmers, M. *Nature* **267**, 694-696 (1977).
13. Terrace, H. S. *Psychol. Sci.* **4**, 162-169 (1993).
14. Cohen, N. J. & Eichenbaum, H. *Memory, Amnesia and the Hippocampal System* (MIT Press, Cambridge, MA, 1993).
15. Tolman, E. C. *Psychol. Rev.* **55**, 189-208 (1948).
16. Dickinson, A. *Contemporary Animal Learning Theory* (Cambridge Univ. Press, 1980).
17. James, W. *The Principles of Psychology* (Holt, New York, 1918).
18. Eichenbaum, H., Otto, T. & Cohen, N. J. *Brain Behav. Sci.* **17**, 449-518 (1994).
19. Gaffan, D. *Expl Brain Res.* **99**, 411-422 (1994).
20. Hunt, M. E., Kesner, R. P. & Evans, R. B. *Psychobiology* **22**, 186-194 (1994).
21. Zola-Morgan, S. M., Squire, L. R. & Ramus, S. J. *Hippocampus* **4**, 483-495 (1994).
22. Nadel, L. *Hippocampus* **5**, 232-234 (1995).
23. Zola-Morgan, S. M., Squire, L. R. & Ramus, S. J. *Hippocampus* **5**, 235-239 (1995).
24. Zola-Morgan, S., Squire, L. R. & Amaral, D. G. *J. Neurosci.* **6**, 2950-2967 (1986).
25. Weiskrantz, L. & Warrington, E. K. in *The Hippocampus Vol. 2* (eds Isaacson, R. L. & Pribram, K. H.) 411-428 (Plenum, New York, 1975).
26. Musen, G. & Squire, L. R. *Neuropsychology* **7**, 119-135 (1993).
27. Eichenbaum, H., Stewart, C. & Morris, R. G. M. *J. Neurosci.* **10**, 331-339 (1990).
28. Eichenbaum, H., Mathews, P. & Cohen, N. J. *Behav. Neurosci.* **103**, 1207-1216 (1989).
29. Kimble, D. P. *Psychol. Bull.* **70**, 285-295 (1968).
30. Douglas, R. J. & Pribram, K. H. *Neuropsychologia* **4**, 197-220 (1966).
31. Jarrard, L. E. *J. Neurosci. Meth.* **29**, 251-259 (1989).
32. Bunsey, M. & Eichenbaum, H. *Hippocampus* (in the press).
33. Moscovitch, M. in *Memory Systems* (eds Schacter, D. L. & Tulving, E.) 269-310 (MIT Press, Cambridge, MA, 1994).

ACKNOWLEDGEMENTS. We thank J. Robisek for assistance in animal testing, J. Dusek for statistical consultation, J. Neidermair for assistance in preparing the histological sections, and D. Amaral, R. Burwell and P. Rapp for consultation on the histological analyses. This work was supported by the NIMH, ONR and NIA.

A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons

Armen N. Akopian, Lucia Sivilotti* & John N. Wood†

Department of Anatomy and Developmental Biology, and * Department of Pharmacology, University College, London WC1E 6BT, UK

DORSAL root ganglion sensory neurons associated with C-fibres, many of which are activated by tissue-damage, express an unusual voltage-gated sodium channel that is resistant to tetrodotoxin¹⁻⁹. We report here that we have identified a 1,957 amino-acid sodium channel in these cells that shows 65% identity with the rat cardiac tetrodotoxin-insensitive sodium channel¹⁰, and is not expressed in other peripheral and central neurons, glia or non-neuronal tissues. *In situ* hybridization shows that the channel is expressed only by small-diameter sensory neurons in neonatal and adult dorsal root and trigeminal ganglia. The channel is resistant to tetrodotoxin when expressed in *Xenopus* oocytes. The electrophysiological and pharmacological properties

†To whom correspondence should be addressed.

a

1	MELPPASVGTTFNRRFTPESLAEIEKQIAAHRAAKKARTKHRGQE..DKG	48	992	SFSADNLTAPEDDGEVNNLQALARIQVLGHRASRAIASYISSHCRFWHP	941
1	MANLLLPRGTSSFRFRFTRESLAAIEKRMAEKQARGGSATSQESREGLQEE	50	944	SFSADNLTAPDEDDGEMNNLQALARIQR.GLRFVKRTTWDFFCCGILRRRP	992
49	EKPRPQLDLKDCNQLPKPYGELPAELVGEPLDLPFYSTHRTFMVLNKS	98	942	KVETQLGMKPPLTSSEAKNHIATDAVSAAVGNLTKPALSSPKENHGDFIT	991
51	EAPRPQLDLQASKKLPDLYGNPPRELIGEPLDLPFYSTQKTFIVLNKG	100	993	KKPAALATHSQLPSCITAPRSPPPPEVEKVPARKETRFEDKRPQGGTP	1042
99	RTISRFSAITWALWLFSPFNLIIRRTAIKVSWSWFISIFITITILVNCVMT	148	992	DPNVVWSVPIAEGESDLDELEEDMEQASQSSWQEDPKPGQQEQLPQVQKC	1041
101	KTIFRFSAITNALYVLSPPHVRRAAVKILVHSLFSLMILMCTILTNCVFMA	150	1043	GDSEPVCVPIAIAESDTEDEEENSLGT..EEESSKQESQVSVSGGHEP	1090
	D-I SI		1042	ENHQAARSAS.MMSSEDLAPYLGESWKRKDSQVPAEGVDDT...SSSE	1087
149	RTDLPEK...VEYVFTVIYTFEALIKILARGFCLNEFTYLRDPWNWLDPS	195	1091	YQEPRAWSQVSETTSSEAGASTSQADWQEQKTEPQAPGCGETPEDSYSE	1140
151	QHDPPWTKYVEYTFATYTFESLVKILARGFCLHAFTFLRDPWNWLDPS	200	1088	GSTVDCPDPEELRKIPELADDDLPDCCFTEGCTRRCPCCNVNTSKSPW	1137
	D-I S2		1141	GSTADMTNTADLLEQIPDLGEDVKDPEDCFTEGCVRRCPCCMVDDTQSPG	1190
196	VITLAYVGAADLRGISGLRTRFVRLRALKTVSVIPGLKVIIVGALIHVSVK	245	1138	ATGWQVRKTCYRIVEHSWFESFIIIFMILLSSGALAFEDNYLEEKPRVKS	1187
201	VIVMAYTTFVLDLGNVSLRTRFVRLRALKTVSISGLKTIIVGALIQSVK	250	1191	KVWVRLRKTCTYRIVEHSWFETFIIFMILLSSGALAFEDIYLEERKTIKVL	1240
	D-I S3 D-I S4			D-III S1	
246	LADVTILTVPCLSVFALVGLQLFKGNLKNKIRNGTDPHKADNLSSEMAE	295	1188	LEYTDRVFTFIFVFEMLLKWVAYGFKKYFTNAWCWLDLIVNISLTSIA	1237
251	LADVMVLTVPCLSVFALIGLQFMGNLRHKCVRNFT.E.LNGTNGSVEADG	299	1241	LEYADKMTFYVFLVLEMLLKWVAYGFKKYFTNAWCWLDLIVDVLVSLVA	1290
	D-I S5			D-III S2 D-III S3	
296	YIF.....IKPGTDPDLLCGNGSDAGHCPGGYVCLKTPDNP	331	1238	KILEYSDVASIKALRTLRLRPLRLRSRFEGRMVVVDALVGAIPSIMNVL	1287
300	LVWNSLDVYLNDPANYLLKNGTTDVLGCGNSSDAGCTCEGYRCLKAGENP	349	1291	NTLGAEMGPIKSLRTLRLRPLRLRSRFEGRMVVVDALVGAIPSIMNVL	1340
				D-III S4	
332	DFNYTSFDSFAWAFSLFLRLMTQD.SWERLYQQTLRASGKMYVFFVLVIF	381	1288	LVCLIFWLIFSIMGVNLFAKFSKCDVTRNPPFSNVNSTMVNKKSECHNQ	1337
350	DHGYTSFDSFAWAFSLFLRLMTQD.CWERLYQQTLRSAGKIYMIFFMLVIF	399	1341	LVCLIFWLIFSIMGVNLFAKFGRCINQTEGDL.P.LNYTIVNKKSECESP	1389
	D-I SS1 D-I SS2			D-III S5	
382	LGSFYLVNLIILAVVTMAYEEQSQATIAIEAKEKKFQAELEVLQKEQEV	431	1338	NSTGHFFVWNVKVNFDNVAMGYLALLQVATFKGMDIMYAAVDSGEINSQ	1387
400	LGSFYLVNLIILAVVAMAYEEQSQATIAETEEKEKRFQAEMLKKEHEAL	449	1390	NVTGELYWTKVNVNFDNVAGYLLALLQVATFKGMDIMYAAVDSRGYEEQ	1439
	D-I S6			D-III SS1 D-III SS2	
432	EALGIDTTSLSQSHSGSPLASKNANERRPRVKSRSVSEGSTD..DNRSPQSD	479	1388	PNWNNLYMYLYFVVFIFFGGFFTLNLFVGVIIIDNFNQKKKLGQDIFM	1437
450	TIRGVDTSRSLSLEMSPLAPVTNHERKSKRRRLSSGTEDGGDRLPKSD	499	1440	PQWEDNLYMYLYFVVFIFFGGFFTLNLFVGVIIIDNFNQKKKLGQDIFM	1489
				D-III S6	
480	PYNQRR...MSF...LGLSSGRRRASGHGSVFHFRAPQDISFPDGITPD	522	1438	TEEQKKKYNNAMKLGSKPKQPIRPLNKYQGFVFDIVTRQAFDIIIMVL	1487
500	SEDGPALNQLSLTHGLSRTSMRPRSSRGSIPTFRRRDQGS..ADFADD	547	1490	TEEQKKKYNNAMKLGSKPKQPIRPLNKYQGFVFDIVTRQAFDVTIMFL	1539
				D-IV SI	
523	DGVFHGDQESRRRG.SILLGR...GAGQTGPLRSPPLQSPNPGRR....	563	1488	ICLNMITMVMVETDEQGEKTKVLGRINQFFVAVFTGECVMKMFALRQYYF	1537
548	ENSTAGESESHRTSLLVWPLRHPSAQGGPGGASAPGYVLNKRNSTVD	597	1540	ICLNMVMTMVMVETDDQSPEKVNILAKINLLFVAIFTGECIVKMAALRHYF	1589
				D-IV S2	
564	.HGEEGLQVPPTGELTAGA.....PEGPALHTTG	591	1538	TNGWNVDFIVVILSIGSLFSAIKSLNENYFSPTLFRVIRLARIGRILR	1587
598	CNGVSVLLGAGDAEATSPGSYLLRPMVLDPRPDPTTTPSEEPGPGQLTPQ	647	1590	TNSWNVDFIVVILSIGSLFSAIKSLNENYFSPTLFRVIRLARIGRILR	1637
				D-IV S3 D-IV S4	
592	QKSFLSAGYLNFPRAQRAMSVVSIMTSVIEELEESKLKCPCLISFAQK	641	1588	LIRAAGKIRTLFALMMSLPALFNIGLLFLVMFIYSIFGMASFANVDE	1637
648	AP...CADGFEEPGARQALSAVSVLTSALAELEESHRKCPPCWNRFAQH	694	1638	LIRGAKIRTLFALMMSLPALFNIGLLFLVMFIYSIFGMASFANVDE	1687
				D-IV S5	
642	YLIWECCPKWRKFKMALFELVTDPPFAELTITLCIVVNTVFMAMEHYPMTD	691	1638	AGIDDMFNFKTPGNSMLCLFQITTSAGWDGLLSPILNTGPPYCDPNLPS	1687
695	YLIWECCPLWMSIKQKVFVMDPPADLTITMCIVLNTLFMALEHYNMTA	744	1688	AGIDDMFNFKTPGNSMLCLFQITTSAGWDGLLSPILNTGPPYCDPNLPS	1737
	D-II SI			D-IV SS1 SS2	
692	AFDAMLQAGNIVFTVFFTMEMAFKIIAFDPYFFQKKWNIFDCVIVTVSL	741	1688	NGSRGNCSPAVGIIFFTTYIIISFLIVNMYIAVILENFVNATEESTEP	1737
745	EFEEMLVQGNLVFTGIFTAEMTFKIIALDPYFFQGGWNIFDSIIIVLSL	794	1738	NGSRGNCSPAVGILFFTTYIIISFLIVNMYIAVILENFVNATEESTEP	1787
	D-II S2 D-II S3			D-IV S6	
742	LELSASKKGLSVLRLTLRLRVFKLAKSWPTLNTLIKIIIGNSVGALGNLT	791	1738	LSEDDFDMFYETWEKFDPEATQFIAPFALSDFADTLGSLRIPKPNQNIL	1787
795	MELGLSRMGNLSVLRFLRLRVFKLAKSWPTLNTLIKIIIGNSVGALGNLT	844	1788	LSEDDFDMFYETWEKFDPEATQFIAYLALSDFADTLGSLRIPKPNQNIL	1837
	D-II S4				
792	FILAIIVFIFALVGKQLLSEYDGCGRKDGVSUWNGEKLRLWHMCDFFHSLV	841	1788	IQMDLPLVPGDKIICLDILFAFTKKNVLGESGELDSLKTINMEEKFMATNLS	1837
845	LVLAIIVFIFAVVGMQLFGKNYSERLHRISD.SGLLPRWHMCDFFHAFLLI	893	1838	INMDLPMVSGDRICMDILFAFTKKNVLGESGEMDALIKIEMEEKFMAANPS	1887
	D-II S5 D-II SS1				
842	VFRILCGEWIENMWVCMEVSQKSICILFLFTVMVLGNLVNLFIALLN	891			
894	IFRILCGEWIENMWVCMEVSQKSICILFLFTVMVLGNLVNLFIALLN	943			
	SS2 D-II S6				

1838 INMDLPMVSGDRIHCMDILFAFTKRVLGESGEMDALKIQMEEKFMAANPS 1887
 1838 KASYEPIATTLRWKQEDLSATVIQAYRSYMLHRS/TLNTHVPAEED 1887
 1888 KISYEPIITTLRRKHEEVATVIQRAFRHLLQRSVKHASFLFRQQAGGS 1937
 1888 GVSLEPG...EGYITFMANSGLPDKSE...TASATSPFPSYDSVTRGL 1928
 1938 GLSDEDAPEREGLIAYMMNGNFSRRSAPLSSSSISSTSPFPSYDSVTRAT 1987
 1929 SDRANINPSSSMQNEDEVAKEGNSPGPQ..... 1957
 1988 SDNLPVRASDYSRSED...LADFPPSPDRDRESIV 2019



FIG. 1 a, Primary sequence of SNS (clone 1B, upper line) aligned with the rat heart tetrodotoxin-insensitive sodium channel¹⁰. The putative transmembrane regions are underlined and domains indicated in subscripts. Sequence gaps are indicated by dots. Conservative substitutions are marked with two dots and non-conservative changes with single dots, and

identical residues are marked with bars. A critical residue implicated in high-affinity TTX binding¹⁴ is in bold-face and underlined. b, *In vitro* transcription/translation (Promega-TNT system) generates a protein of size ~220K (marked) predicted by the 1,957 amino-acid open reading frame in a. Track A, SNS clone; B, control (no plasmid DNA).

METHODS. DRG from all spinal levels of neonatal Sprague-Dawley rats were isolated, and RNA extracted using guanidinium isothiocyanate and acidic phenol extraction¹³. Poly(A)⁺ RNA was isolated by chromatography on oligo-dT cellulose and used to generate a directionally cloned oligo-dT primed rat DRG cDNA library. First-strand cDNA was generated by reverse transcription (Superscript BRL) using oligo-dT-Xho-primer adaptors, and methylated dNTPs supplied in the Stratagene directional cloning Zap-II kit. Second-strand DNA was synthesized using DNA polymerase I, RNase H and *E. coli* ligase (BRL). Blunt-ended (T4 polymerase) cDNA was ligated to *EcoRI* adaptors. The > 6 kb fraction of *EcoRI*-*XhoI* cDNA was isolated by gel electrophoresis and ligated into lambda zap II (Stratagene synthesis kit). The size-fractionated library was screened by hybridization with a ³²P dATP random-primed (Gibco Rad-prime kit) SNS-like *SphI* fragment probe (position 6,088–6,453) of the 3'UTR of the cDNA¹³. The filters were given a final wash in 0.2 × SSC, 0.5% SDS at 65 °C for 30 min. 20 positive clones were isolated from 2 × 10⁵ PFU of which 5 were either full-length or contained the initiator methionine. The region upstream of the translational initiation site was identical to that of the cardiac sodium channel 5'UTR, and contained a Kozak consensus sequence (tgaggaagATGG). Of the 20 positive clones, 2 contained an insertion in the interdomain region I (one of these was full-length). A further clone had an extended 3'UTR of ~2 kb that appears to be the result of using an alternative polyadenylation site rather than alternative splicing in the 3'UTR. Two clones (1B and 1C) were completely sequenced on both strands using the dideoxy method. The clones showed essentially identical sequence except for 8 differences, 5 of which resulted in altered amino-acid sequence F16–S16, L393–P393, T470–I470, R278–H278, I1,876–M1,876. ³⁵S-methionine-labelled proteins were solubilized in 4% SDS and separated on 10% SDS–PAGE in the presence of reducing agents, autoradiographed with Kodak X-omat film, and *M_s* determined with reference to rainbow molecular mass markers (Amersham).

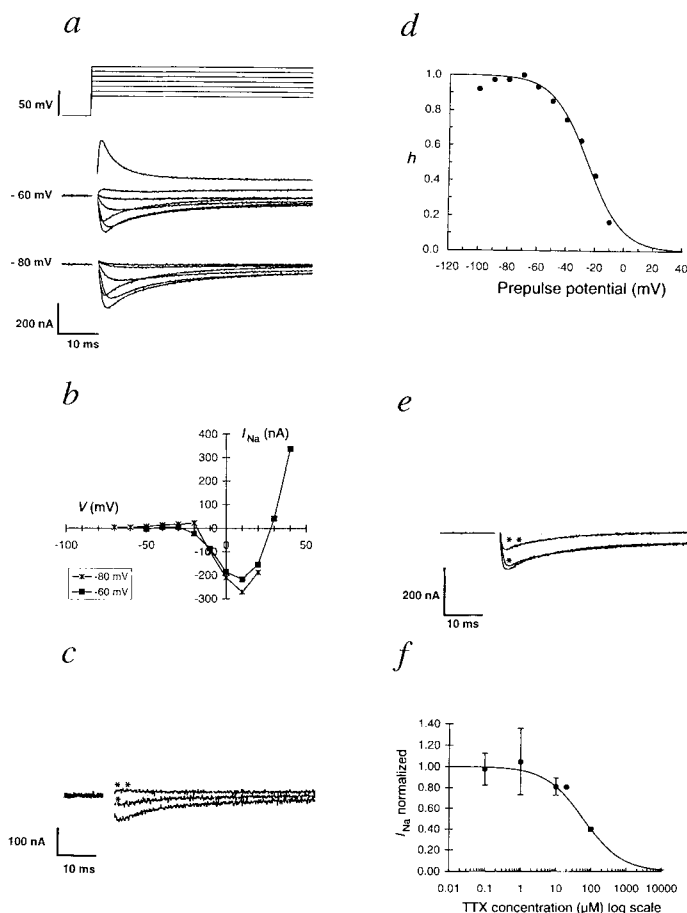


FIG. 2 Electrophysiological properties of SNS expressed in *Xenopus* oocytes. a, Inward currents evoked by depolarizing steps from -60 mV to a command potential of -20 to +40 mV (in 10 mV steps, top) and from -80 mV to a command potential of -30 to +20 mV (bottom) in oocytes injected with SNS cRNA. Current traces are blanked for the first 1.5 ms from the onset of the voltage step to delete the capacity transients for clarity. b, Current-voltage curves from the traces shown in a. Note that the peak current is reached at the same command voltage for the two holding potentials, but is slightly smaller from -60 mV because of steady-state inactivation. c, Effect of 50% (*) or 100% (**) replacement of external Na⁺ by N-methyl-D-glucosamine on the SNS current elicited by stepping the oocyte from -60 to +10 mV. d, Steady-state voltage-dependence of inactivation for the SNS Na⁺ current in a representative oocyte (prepulse duration 1 s, midpulse 5 ms). Data were fitted with the equation $h_{\infty} = 1/(1 + \exp((V - V_{50})/k))$, where *V* is the prepulse potential, *V*₅₀ the potential of 50% inactivation and *k* the slope factor (least-squares fit): in this oocyte *V*₅₀ = -25.4 mV, *k* = 12.1 mV. e, Effect of TTX (*10 μM and **100 μM) on the peak Na⁺ current (test pulse from -60 to +20 mV). The effect was quickly reversible upon washout. f, TTX dose-response curve, fitted with the Hill equation; Hill slope = 0.81 ± 0.23, *n* = 3 (CVFIT, courtesy of D. Colquhoun).

METHODS. SNS-1B encoding pBluescript SK plasmid DNA was digested to position -21 upstream of the initiator methionine (Promega Erase-a-base system). The linearized and digested plasmid was cut with *KpnI* and subcloned into an oocyte expression vector pSp64GL (*Sma*-*KpnI*) sites. pSp64GL is derived from pSp64.T. pSp64.T was cut with *SmaI*-*EcoRI*, blunt-ended with Klenow enzyme, and recircularized. Part of the pGEM7Zf (+) polylinker (*SmaI*-*KpnI*-*EcoRI*-*XhoI*) was ligated into the blunt-ended *BglII* site of pSp64.T. This vector, with an altered polylinker for DNA inserts (*SmaI*-*KpnI*-*EcoRI*-*XhoI*), and linearization (*Sall*-*XbaI*-*BamHI*) was named pSp64GL. The resulting plasmid was linearized with *XbaI*, and cRNA transcribed with SP6 polymerase using 1 mM 7-methylGppG²². cRNA (70 ng) was injected into *Xenopus* oocytes 7–14 days before recording: immature, stage IV oocytes were chosen because of their smaller diameter and therefore capacitance²⁷. Oocytes were impaled with 3 M KCl electrodes (< 1 MΩ) and perfused at 3–4 ml min⁻¹ with modified Ringer solution containing (in mM) NaCl 115, KCl 2.5, HEPES 10, MgCl₂ 1.8, CaCl₂ 1, pH 7.2, at a temperature of 19.5–20.5 °C. Digital leak subtraction of two-electrode voltage-clamp current records was carried out using as leak currents produced by hyperpolarizing pulses of the same amplitude as the test depolarizing commands. Oocytes in which leak commands elicited time-dependent currents were discarded. Averages of 10 records were used for both test and leak.

of the expressed channel are similar to those described for the small-diameter sensory neuron tetrodotoxin-resistant sodium channels. As some noxious input into the spinal cord is resistant to tetrodotoxin^{8,11,12}, block of expression or function of such a C-fibre-restricted sodium channel may have a selective analgesic effect.

We have identified a 6.5 kilobase (kb) transcript expressed

selectively in dorsal root ganglia (DRG), that shows sequence similarities with known voltage-gated sodium channels¹³. Sequencing the full-length complementary DNA clone revealed an open reading frame of 5,871 nucleotides encoding a 1,957 amino-acid protein, named SNS, that showed 65% identity at the amino-acid level with the rat cardiac tetrodotoxin-insensitive (TTXi) sodium channel (Fig. 1a)¹⁰. Interestingly, the aromatic residue that is

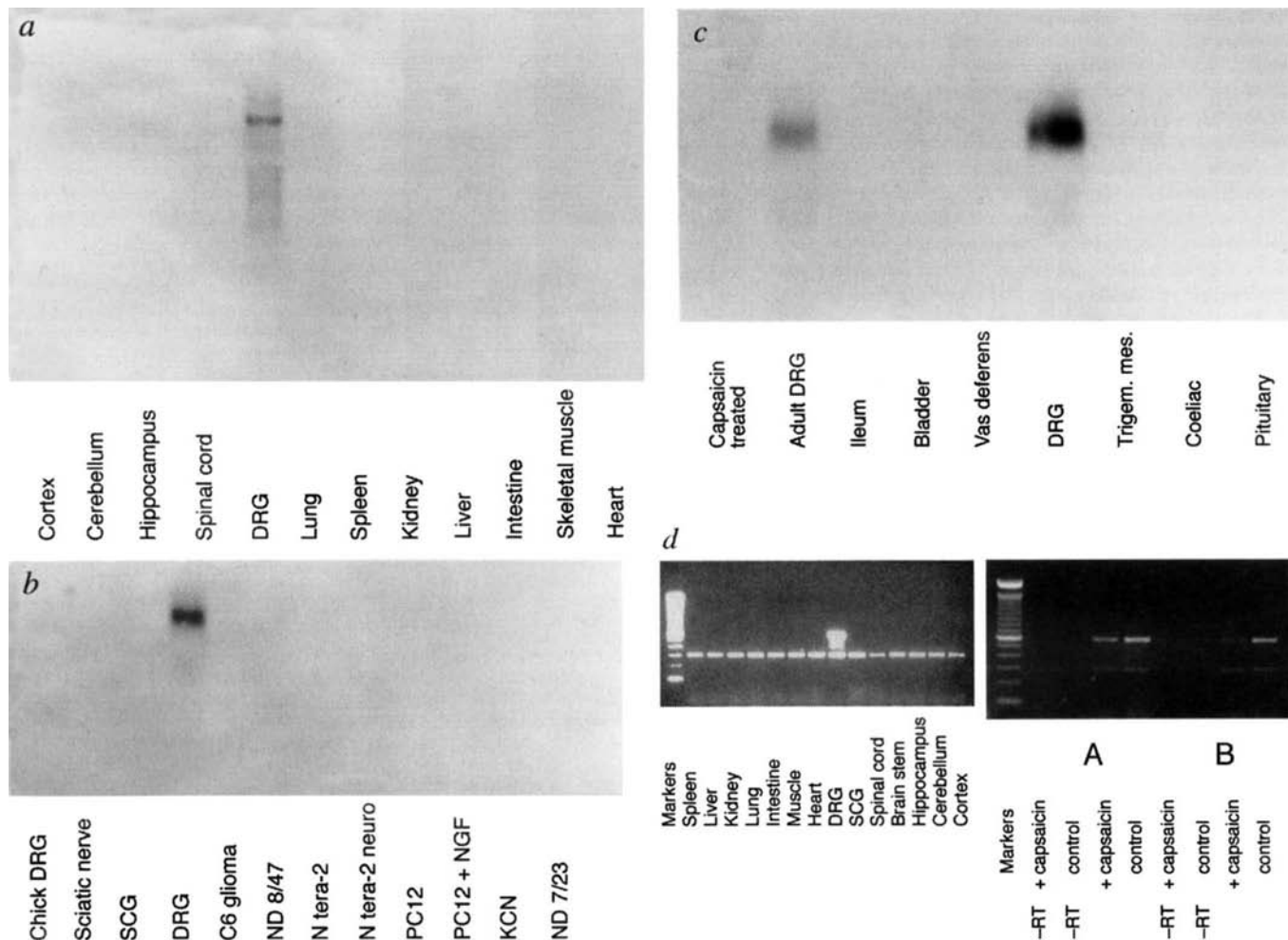


FIG. 3 Northern blot of SNS distribution in neonatal and adult rat tissues and cell lines. *a*, Central nervous system and non-neuronal tissues from neonatal rats. *b*, Peripheral nervous tissue including neonatal Schwann cells and sympathetic neurons, as well as C6 glioma, human embryonal carcinoma line N-tera-2 and N-tera-2 neuro, rat sensory neuron-derived lines ND7 and ND8, and human neuroblastoma SMS-KCN and PC12 cells grown in the presence of NGF²². *c*, Adult rat tissue with neonatal DRG control (DRG), including pituitary, superior cervical ganglia, coeliac ganglia, trigeminal mesencephalic nucleus, vas deferens, bladder, ileum and DRG of adult animals treated with capsaicin (50 mg kg⁻¹) at birth. Total RNA (10 µg, *a*), or 25 µg (*b* and *c*) apart from superior cervical ganglion sample (*b*, 10 µg) and capsaicin-treated adult rat DRG (5 µg) were northern blotted. Lane labels are beneath each panel. *d*, RT-PCR confirmation of northern data shown using ethidium bromide stained agarose gels. Left, RT-PCR of oligo dT-primed cDNA from various tissues using SNS primers (top band) and L-27 ribosomal protein primers (lower band)²⁸ showing the presence of SNS transcripts in DRG tissue only. Right, RT-PCR of SNS and L-27 transcripts from DRG neurons cultured and treated with capsaicin (overnight 10 µM) (A) or dissected from neonatal animals treated with capsaicin (50 mg kg⁻¹) on 2 consecutive days, followed by DRG isolation 5 days later; B). Note the identical signal for the L-27 probe and the significant diminution in SNS signal from capsaicin-treated cultures or animals. Control PCR reactions without reverse transcriptase treatment are also shown, to control for contaminating genomic DNA.

METHODS. Random primed ³²P-labelled DNA Pst-Acc1 fragment probes (50 ng, specific activity 2 × 10⁹ c.p.m. per µg DNA) from interdomain

region I (nucleotide position 1,478–1,892) were used to probe total RNA extracted from tissues¹³. Total RNA was separated on 1.2% agarose-formaldehyde gels, and capillary blotted onto Hibond-N (Amersham)²⁹. The amounts of RNA on the blot were roughly equivalent, as judged by ethidium bromide staining of ribosomal RNA and by hybridization with the ubiquitously expressed L-27 ribosomal protein transcripts^{28,29}. Filters were prehybridized in 50% formamide, 5 × SSC containing 0.5% SDS, 5 × Denhardt's solution, 100 µg ml⁻¹ boiled sonicated salmon sperm DNA (average size 300 bp), 10 µg ml⁻¹ poly-U and 10 µg ml⁻¹ poly-C at 45 °C for 6 h. After 36 h hybridization in the same conditions using 10⁷ c.p.m. per ml hybridization probe, the filters were briefly washed in 2 × SSC at room temperature, then twice with 2 × SSC with 0.5% SDS at 68 °C for 15 min, followed by a 20 min wash in 0.5% SDS, 0.2 × SSC at 68 °C. The filters were autoradiographed overnight (*a*) or for 4 days (*b* and *c*) on Kodak X-omat film on the blots shown here. For RT-PCR experiments, 10 µg total RNA from neonatal rat tissues (*a*), or 2 µg total RNA from control or capsaicin-treated rat DRG or DRG neurons in culture (*b*) were treated with DNase I and extracted with acidic phenol to remove genomic DNA. cDNA was synthesized with Superscript reverse transcriptase using oligo dT(12–18) primers and purified on Qiagen 5 tips. PCR was used to amplify cDNA (35 cycles, 94 °C, 1 min; 55 °C, 1 min; and 72 °C, 1 min), and products separated on agarose gels before staining with ethidium bromide. L-27 primers²⁹ were added to the PCR reaction 5 cycles after the start of the reaction with SNS-specific primers which comprised 5'-CAGCTTCGCTCAGAAGTATCT-3', 5'-TTCTCGCCGTTCCACACGGAGA-3'.

involved in high-affinity binding of TTX to the channel atrium of TTX-sensitive sodium channels¹⁴ was altered to a hydrophilic serine in the predicted SNS protein, whereas the residues implicated in sodium-selective permeability were conserved¹⁵. In order to test whether the putative protein was encoded by the cDNA clone, we used a coupled *in vitro* transcription/translation system, and found that a protein of the predicted M_r (220.5K) was encoded by the open reading frame (Fig. 1b).

DRG neurons express at least three types of sodium channels which differ in kinetics and sensitivity to TTX^{2,4,7}. Neurons with small-diameter cell bodies and unmyelinated axons (C-fibres) include most of the nociceptor (damage-sensing) population and express both a fast TTX-sensitive current, and a slower TTXi current^{3,4}. Of the five cloned sodium channel α -subunit transcripts known to be present in DRG¹⁶⁻¹⁹ none exhibits the properties of the TTXi channel.

We tested the functional properties of the protein encoded by the SNS clone by expression in *Xenopus* oocytes. After a minimum incubation of 7 days from cRNA injection, step depolarizations to potentials positive to -30 mV elicited inward currents (Fig. 2a) which peaked between $+10$ and $+20$ mV with an average maximum amplitude of 164 ± 72 nA (from -60 mV holding potential, $n = 13$) and a reversal potential of $+35.5 \pm 2.2$ mV ($n = 10$; Fig. 2b). The low level of expression precluded the use of the patch-clamp techniques necessary for detailed kinetic analysis in oocytes¹⁵. Nevertheless, the gross properties of the recombinant channel could be observed. The inward current was reversed by total replacement of Na^+ in the external medium with an impermeant cation (Fig. 2c); its reversal potential was shifted in 50% Na^+ by 13.7 ± 3.2 mV in the hyperpolarizing direction ($n = 3$; predicted value for a Na^+ -selective channel, 17.5 mV). The inactivation produced by a 1 s prepulse was half-maximal at -30.0 ± 1.3 mV (slope factor 14.0 ± 1.7 mV, $n = 5$; Fig. 2d). TTX had no effect at nanomolar concentrations, and produced only a $19.1 \pm 8.3\%$ reduction at $10 \mu\text{M}$ (Fig. 2e, $n = 3$). The estimated half-maximal inhibitory concentration (IC_{50}) was

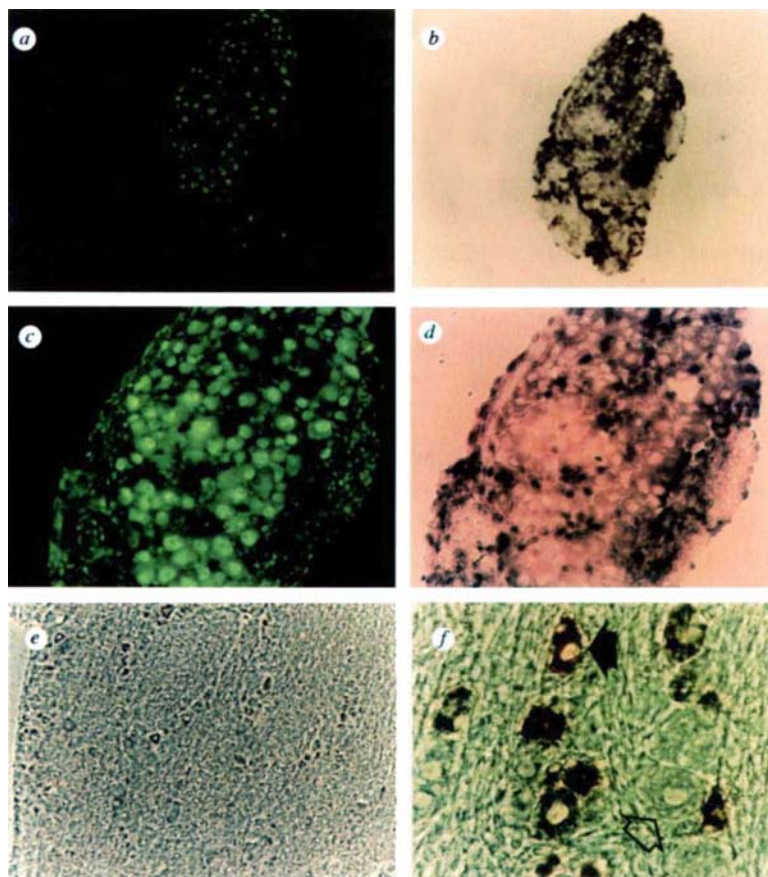
$59.6 \pm 10.1 \mu\text{M}$. As in small-diameter DRG neurons³, the local anaesthetic lignocaine was also weakly inhibitory, producing a maximum block of $41.7 \pm 5.4\%$ at 1 mM on the peak current elicited by depolarizing pulses from -60 mV to $+10$ mV (1 every min; $n = 3$), whereas under the same conditions $100 \mu\text{M}$ phenytoin had no effect. Another similarity with the TTXi Na^+ current of DRG neurons was the effectiveness and rank order of Pb^{2+} versus Cd^{2+} in reducing peak Na^+ currents ($-63.9 \pm 18.1\%$ for Pb^{2+} versus $-24.4 \pm 7.9\%$ for Cd^{2+} at $50 \mu\text{M}$ and $100 \mu\text{M}$, respectively; $n = 3$, $P = 0.0189$, randomization test RANTEST). The electrophysiological and pharmacological characteristics of the oocyte-expressed SNS channel are thus similar to the properties of the sensory neuron TTXi channel³, given the constraints of expression in an oocyte system. In oocytes expressing SNS, the peak of the I/V plot occurred at a more depolarized potential than that of the DRG TTXi current, despite a similar reversal potential³. This difference may reflect the absence of the accessory $\beta 1$ subunit found in DRG²⁰, which is known to shift activation to more negative potentials when expressed with the subunit of other Na^+ channels²¹.

Because the properties of the channel are similar to those observed in small-diameter sensory neurons, we examined the tissue and cell-type distribution of the transcript. We used northern blots and *in situ* hybridization to examine neonatal and adult rat tissues for the expression of SNS messenger RNA. We could not detect the transcript in any non-neuronal tissues or in the central nervous system using northern blots or reverse transcription of mRNA and the polymerase chain reaction (RT-PCR) (Fig. 3).

Sympathetic neurons from the superior cervical ganglion and Schwann cell-containing sciatic nerve preparations, as well as several neuronal cell lines were also negative. However, total RNA extracts from neonatal and adult rat DRG gave a strong signal of size about 7 kb on northern blots (Fig. 3). These data suggest that the SNS channel is not expressed only in early development²⁻⁹. When neonatal rats were treated with capsaicin

FIG. 4 *In situ* hybridization carried out using digoxigenin-labelled cRNA transcripts from unique SNS sequence. a, Fluorescence micrograph ($\times 40$) of a $7\text{-}\mu\text{m}$ section of neonatal rat lumbar DRG stained with FITC-anti-mouse Fab2 antibody detected monoclonal antibody RT97 directed against neurofilaments²². b, *In situ* hybridization of the same section as a with anti-SNS complementary RNA. c, d, Higher power pictures of a and b ($\times 100$) showing mutually exclusive patterns of expression of RT97 neurofilament-immunoreactivity (some nuclear staining is apparent) and SNS mRNA expression. e, Control (sense) *in situ* hybridization of adult trigeminal ganglion ($\times 200$). f, Single-cell resolution of adult trigeminal ganglion neurons specifically hybridized with SNS probe ($\times 400$). Note the clearly negative large diameter cell body neurons (light arrow) and the positively stained small diameter neurons (dark arrow).

METHODS. An SNS PCR fragment of interdomain region I between positions 1,736 and 1,797 was subcloned into pGem3Z, and DIG-UTP labelled sense or antisense cRNA generated using SP6 or T7 polymerase, respectively. Sample preparation, hybridization and visualization of *in situ* hybridization with alkaline phosphatase conjugated anti-DIG antibodies was carried out as described in ref. 30 with modifications. Frozen tissue sections ($10\text{-}\mu\text{m}$ -thick) were fixed for 10 min in PBS containing 4% paraformaldehyde. Sections were acetylated in 0.1M triethanolamine, 0.25% acetic anhydride for 10 min. Prehybridization was carried out in 50% formamide, $4 \times \text{SSC}$, $100 \mu\text{g ml}^{-1}$ boiled and sonicated ssDNA, $50 \mu\text{g ml}^{-1}$ yeast tRNA, $2 \times \text{Denhardt's}$ solution at room temperature for 1 h; hybridization was carried out overnight in the same buffer at 65°C . Probe concentration was 50 ng ml^{-1} . Sections were washed in $2 \times \text{SSC}$ at 72°C for 1 h and twice in $0.1 \times \text{SSC}$ for 30 min at 72°C , before visualization at room temperature with anti-digoxigenin alkaline phosphatase conjugated antibodies. The same sections were then stained with mouse monoclonal antibody RT97²².



and total adult DRG RNA subsequently examined by northern blotting, the signal was substantially reduced, suggesting that the SNS transcript is expressed selectively by capsaicin-sensitive (predominantly nociceptive) neurons. These data were confirmed by RT-PCR experiments on both cultures of DRG neurons, and in whole animal studies (Fig. 3). We used *in situ* hybridization to examine the expression of SNS transcripts at the single-cell level in both adult trigeminal ganglia and neonatal and adult rat DRG (Fig. 4).

Subsets of sensory neurons from both tissues showed intense signals with an SNS-specific probe. Combined immunohistochemistry with the large-diameter neuron-specific monoclonal antibody RT97, showed that most of these cells did not express the SNS transcript (Fig. 4), confirming an association between small-diameter C-fibre-associated capsaicin-sensitive neurons and expression of SNS²².

It therefore seems likely that SNS accounts for the TTXi

sodium currents described in sensory neurons in culture³, and may play a role in nociceptive transmission because some noxious input to the central nervous system is known to be insensitive to TTX^{8,11,12}. Persistent activation of peripheral nociceptors has been found to result in changes in excitability in the dorsal horn associated with the establishment of chronic pain²³. Increased sodium channel activity has also been shown to underlie neuroma-induced spontaneous action potential generation²⁴. Conversely, chronic pain may be successfully treated by surgical or pharmacological procedures which block peripheral nerve activation^{25,26}. Blockade of nociceptor input may therefore produce useful therapeutic effects, even though central nervous system plasticity plays a pivotal role in the establishment of chronic pain. Sensory neuron-specific voltage-gated sodium channels, particularly subtypes associated with a nociceptive modality such as SNS, thus provide potential targets for therapeutic intervention in a range of pain states. □

Received 15 August; accepted 9 November 1995.

- Matsuda, Y., Yoshida, S. & Yonezawa, T. *Brain Res.* **154**, 69–82 (1978).
- Kostyuk, P. G., Veselovsky, N. S. & Tsyndrenko, A. Y. *Neuroscience* **6**, 2423–2430 (1981).
- Roy, M. L. & Narahashi, T. J. *Neurosci.* **12**, 2104–2111 (1992).
- Caffrey, B., Eng, D., Black, J., Waxman, S. G. & Kocsis, J. *Brain Res.* **592**, 283–297 (1992).
- Cohen, S. A. & Barchi, R. L. *Int. Rev. Cytol.* **137c**, 55–103 (1993).
- Nowycky, M. in *Sensory Neurons* (ed. Scott, S.) 97–116 (Oxford Univ. Press, Oxford, 1993).
- Elliott, A. A. & Elliott, J. R. *J. Physiol.* **463**, 39–56 (1993).
- Jeftinija, S. *Brain Res.* **639**, 125–134 (1994).
- Carrier, G. O. & Ikeda, S. R. *Pflügers Arch.* **421**, 7–16 (1992).
- Rogart, R. B., Cribbs, L. L., Muglia, L. K., Kephart, D. D. & Kaiser, M. W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8170–8174 (1989).
- Jeftinija, S. *Brain Res.* **665**, 69–76 (1994).
- Del-Bianco, E., Maggi, C. A., Santicoli, P. & Geppetti, P. *Neurosci. Lett.* **170**, 163–166 (1994).
- Akopian, A. N. & Wood, J. N. *J. biol. Chem.* **270**, 21264–21270 (1995).
- Satin, J. et al. *Science* **256**, 1205–1207 (1992).
- Heinemann, S., Terlau, H., Stühmer, W., Imoto, K. & Numa, S. *Nature* **356**, 441–443 (1992).
- Beckh, S. *FEBS Lett.* **262**, 317–322 (1990).
- Waxman, S. G., Kocsis, J. D. & Black, J. A. *J. Neurophysiol.* **72**, 466–470 (1994).
- Schaller, K. L., Krzymien, D. M., Yarwosky, P. J., Kruger, B. K. & Caldwell, J. H. *J. Neurosci.* **15**, 3231–3242 (1995).
- Gautron, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7272–7276 (1992).

- Oh, Y., Sashihara, S. & Waxman, S. G. *Neurosci. Lett.* **176**, 119–122 (1994).
- Mandel, G. *Curr. Opin. Neurobiol.* **3**, 278–282. (1993).
- Chen, C. et al. *Nature* **377**, 428–432 (1995).
- Willis, W. D. in *Proc. 7th World Cong. Pain* (eds Gebhart, G. F., Hammond, D. L. & Jensen, T. S.), 301–325 (International Association for the Study of Pain, Seattle, 1994).
- Matzner, O. & Devor, M. *J. Neurophysiol.* **72**, 349–359 (1994).
- Chabal, C., Russell, L. C. & Burchiel, K. J. *Pain* **38**, 333–338 (1989).
- Raj, P. P. in *Neural Blockade in Clinical Anaesthesia and Management* (eds Cousins, M. J. & Bridenbough, J. B.) 899–933 (Lippincott, Philadelphia, 1988).
- Dumont, J. N. *J. Morphol.* **136**, 153–180 (1972).
- LeBeau, M.-C., Alvarez-Bolando, G., Braissant, O., Wahli, W. & Catsicas, S. *Nucleic Acids Res.* **19**, 1337 (1991).
- Ninkina, N. N., Stevens, G., Wood, J. N. & Richardson, W. D. *Nucleic Acids Res.* **21**, 3175–3182 (1993).
- Schaeren-Wimers, N. & Gerfin-Moser, A. *Histochemistry* **100**, 431–440 (1993).

ACKNOWLEDGEMENTS. We thank D. Colquhoun for advice and support, N. Abson, D. Attwell, C.-C. Chen, R. Docherty, A. Dolphin, O. Krylova and N. Orlie for critical advice; E. Campbell, Y. Vallis and A. Dickenson for help with dissections, L. Dale for pSP64T, P.W. Andrews for N-tera-2 and S. McMahon and M. Fitzgerald for capsaicin-treated rat tissues. Electrophysiology experiments were performed with WCP software, courtesy of J. Dempster, Strathclyde University. We are very grateful to the Wellcome Trust and the Medical Research Council who supported this work. SNS has the EMBL accession number X 92184.

Chondrodysplasia and neurological abnormalities in ATF-2-deficient mice

Andreas M. Reimold*†, Michael J. Grusby*†, Bela Kosaras‡, Jochen W. U. Fries§, Ryuji Mori||, Sokichi Maniwa||, Isabelle M. Clauss*, Tucker Collins§, Richard L. Sidman‡, Melvin J. Glimcher|| & Laurie H. Glimcher*†

* Department of Cancer Biology, Harvard School of Public Health,

† Department of Medicine, Harvard Medical School,

‡ Division of Neurogenetics, New England Regional Primate Research Center, Harvard Medical School, § Vascular Research Division, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, and || Department of Orthopaedic Surgery, Laboratory for the Study of Skeletal Disorders and Rehabilitation, Harvard Medical School, Children's Hospital, Boston, Massachusetts 02115, USA

ACTIVATING transcription factor-2 (ATF-2) is a basic region leucine zipper protein whose DNA target sequence is the widely distributed cAMP response element (CRE). We report here that mice carrying a germline mutation in ATF-2 demonstrated unique actions of ATF-2 not duplicated by other ATF/CREB family members. Mutant mice had decreased postnatal viability and growth, with a defect in endochondral ossification at epiphyseal plates similar to human hypochondroplasia. The

animals had ataxic gait, hyperactivity and decreased hearing. In the brain, there were reduced numbers of cerebellar Purkinje cells, atrophic vestibular sense organs and enlarged ventricles. Unlike CREB α/δ -deficient mice whose main defect is in long-term potentiation, the widespread abnormalities in ATF-2 mutant mice demonstrate its absolute requirement for skeletal and central nervous system development, and for maximal induction of select genes with CRE sites, such as E-selectin.

A mutation of the ATF-2 gene was introduced into the germ line of mice (Fig. 1a–c) and confirmed by northern blotting, Southern blotting of reverse-transcribed polymerase chain reaction (RT-PCR) products (Fig. 1d,e), western blotting of brain extract (Fig. 1f), and mobility shift assay using ATF-2 antibody (Fig. 1g).

Survival was markedly affected by disruption of the ATF-2 gene (Fig. 2g). Only 51% of homozygous (–/–) mice survived past 1 month. Two waves of deaths occurred: within the initial 2 days of life and near weaning. The lifespan of surviving (–/–) mice was, however, normal. Heterozygous (+/–) mice had a 50% decrease in ATF-2 transcripts in brain (Fig. 1e) with an increased mortality rate roughly half that of –/– mice, consistent with a gene dosage effect (Fig. 2g). Surviving –/– mice were mildly to severely runted, reaching, on average, 71% of the weight of sex-matched (+/+) littermates. Half the –/– animals spontaneously developed periocular neutrophilic/lymphocytic infiltrates (data not shown).

Potential ATF-2 binding sites are found in promoter regions of many genes involved in skeletal development, including osteopontin, osteocalcin and alkaline phosphatase^{1–4}. In developing bone, ATF-2 messenger RNA is found in resting and proliferating, but not hypertrophic, chondrocytes in the cartilaginous

portions of ribs, vertebrae and long bones, as seen by *in situ* hybridization of wild-type mouse embryos (Fig. 2b,c). ATF-2 $-/-$ mice showed uniform dwarfism, with the reduced spine length proportional to the shortened extremities (Fig. 2a). The limb length decrease was rhizomelic, with more severe shortening in proximal bony segments than distal ones. Consistent with the distribution of ATF-2 in cartilage, the main abnormality found in the bones of $-/-$ animals was disorganization in the sequence of endochondral ossification at epiphyseal growth plates. Histology showed abnormal clusters of cells instead of ordered columns of chondrocytes, seams of bone covering the metaphyseal side of the physes, and lack of ingrowth of chondroclasts and capillaries (Fig. 2d,e). This resulted in the absence of normal-appearing cartilaginous trabeculae extending from the physes and led to distinctly thin epiphyseal growth plates. *In vivo* labelling of cartilage cells with ^3H -thymidine confirmed that cartilage cell division was dramatically reduced (Fig. 2f). The histological appearance of the epiphyseal plates coupled with the variable penetrance of the bony defect is similar to human hypochondroplasia. Hypochondroplasia and achondroplasia may be allelic disorders⁵ and the gene mutated in achondroplasia is the fibroblast growth factor receptor-3 (FGFR3)^{6,7}. However, FGFR3 mRNA levels are normal in ATF-2 $-/-$ mice.

ATF-2 mRNA is expressed preferentially in brain^{8,9}. In ATF-2 $-/-$ mice, brain histology showed enlargement of the ventricles,

reduction of brain size, and distinctive cerebellar defects. At the cerebellar molecular-granular cell layer boundary, a 50% decrease in Purkinje cells was found (Fig. 3a,b), representing an absolute deficit of neurons, because no other neuron cell type occupied the gaps between Purkinje cell bodies. Simultaneously, numerous large neurons having the size and shape of Purkinje cells were distributed throughout the granular cell layer. Partially correlating with the cerebellar abnormalities, $-/-$ mice showed a characteristic fine-amplitude whole-body tremor, twitches of the head, and decreased hind limb coordination. In the inner ear vestibular system, the maculae of both the saccule and utricle were markedly atrophic, with significant reductions in numbers of sensory cells, stereocilia and otoconia (Fig. 3c,d), whereas nearby vestibular ganglia had decreased neuron cell bodies (Fig. 3e,f). ATF-2 $-/-$ animals displayed the triad of decreased hearing, hyperactivity and vertical head tossing, which is a common result of inner ear pathology in mice. The most striking feature of the brain histology is the presence of Purkinje-like cells in the otherwise normal-appearing granular cell layer of the cerebellum, a finding not observed in any other known mouse mutant. The abnormal placement of these cells suggests that ATF-2 may regulate the migration of these specialized neurons to their proper location.

To assess the *in vivo* effect of the ATF-2 mutation on a promoter with a well characterized ATF-2 binding site, E-selectin

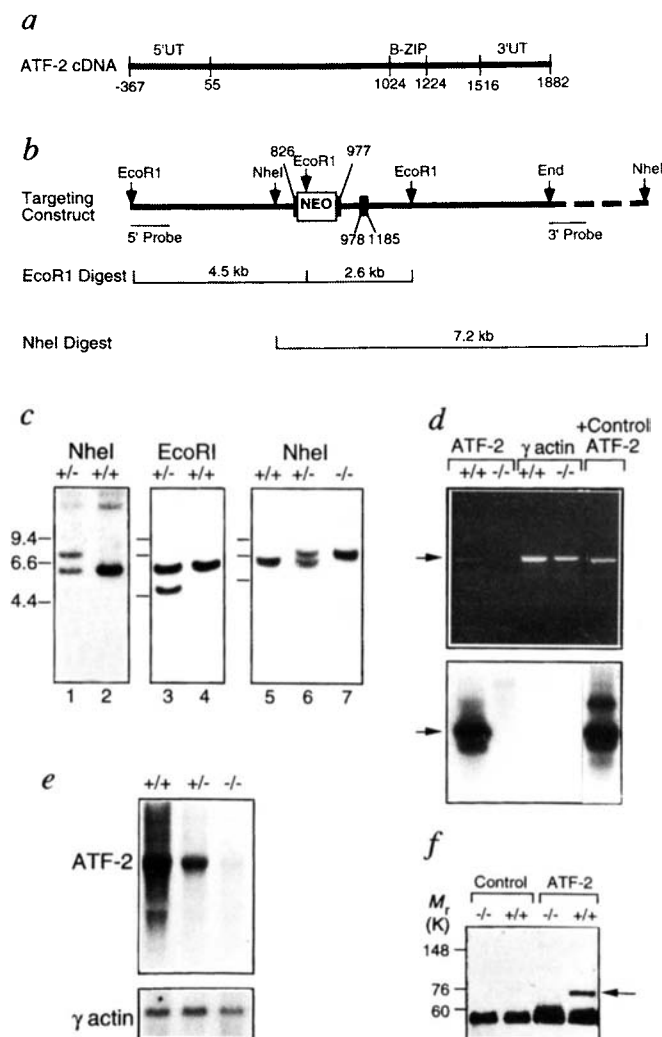


FIG. 1 Targeted disruption of the ATF-2 gene. a, The murine ATF-2 cDNA. b, The 9 kb ATF-2 targeting construct extends from the 5' EcoRI site to the position marked 'End' and includes two exons (cDNA positions 826–1,185, numbering as in ref. 24). Insertion of *neo* at cDNA position 899 added a new

EcoRI site. c, Mutated ATF-2 was detected by Southern blotting of genomic DNA from ES cells (lanes 1–4) or (129/Sv $\times$ BALB/c) F_2 tail biopsies (lanes 5–7), digested with EcoRI and hybridized with the 5' probe (lanes 3, 4) or digested with NheI and hybridized with the 3' probe (lanes 1, 2 and 5–7). d, RT-PCR of ATF-2 $+/+$ and $-/-$ RNA. An ethidium-stained gel (top) shows PCR products generated by ATF-2-specific or γ -actin primers. The control lane contains ATF-2 cDNA. A Southern blot of the same gel (bottom) was probed with ATF-2 cDNA sequences 736–967. e, Northern blot of mRNA from ATF-2 $+/+$, $+/-$ or $-/-$ brain hybridized with ATF-2 cDNA sequences 736–967 (top) or γ -actin (bottom). f, Western blot of brain extract from $+/+$ or $-/-$ mice immunoprecipitated with control mouse IgG or with anti-ATF-2 monoclonal antibody (Santa Cruz Biotechnology). g, Mobility shift assay of $+/+$ and $-/-$ nuclear extract using a CRE probe and antibody to mouse IgG (–lanes) or ATF-2. h, Northern blotting of mRNA from $+/+$ or $-/-$ lung induced with LPS 150 μg per mouse and probed with *c-jun* or *c-fos*, or from brain probed with CREB.

METHODS. The targeting construct, prepared using an ATF-2 genomic clone obtained from a BALB/c cosmid library, was electroporated into ES-D3 cells, and selected with Geneticin and gancyclovir²⁵. Homologously recombined clones (frequency of 17%) were injected into BALB/c blastocysts and one clone resulted in germline transmission. Immunoprecipitation and western blotting were performed as described²⁶. Nuclear extracts were prepared from potato-starch-induced peritoneal macrophages and used in mobility shift assays as described¹³.

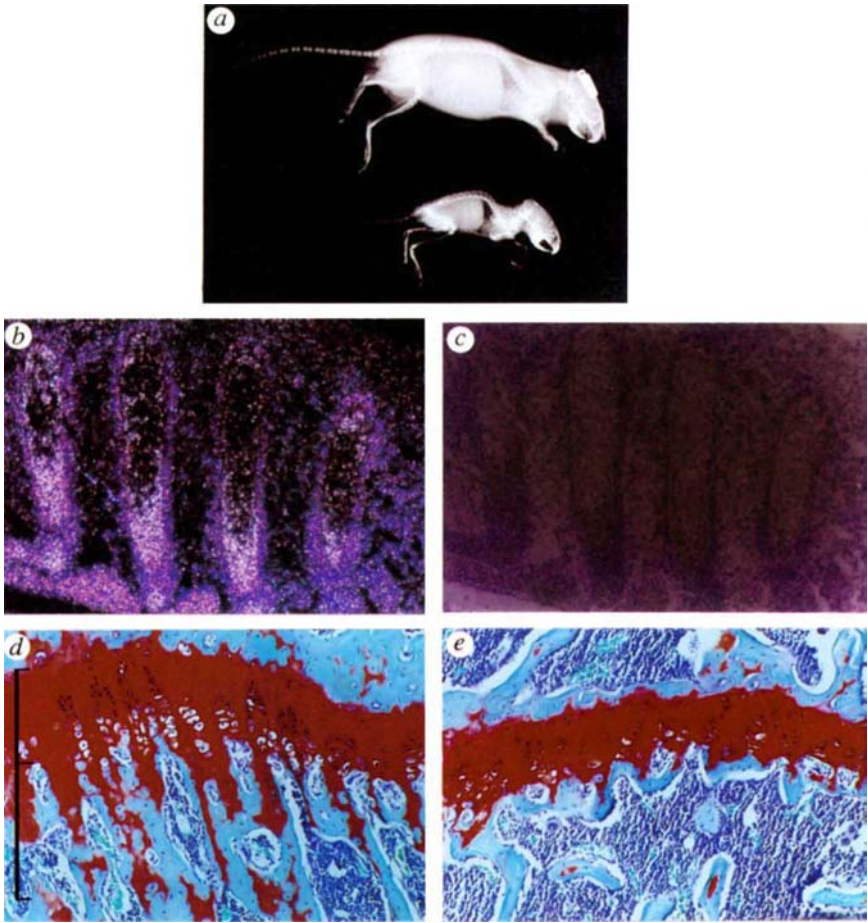
gene induction was studied¹⁰⁻¹². Intraperitoneal injection of lipopolysaccharide (LPS) at doses up to 50 µg resulted in readily detectable E-selectin protein (Fig. 4) and mRNA (not shown) in control animals, but reduced or absent induction in -/- mice. The control protein VCAM-1, whose promoter lacks ATF-2 sites, was induced in both -/- and +/+ tissues. Higher doses of LPS led to equal induction of E-selectin mRNA in +/+ and -/- animals. Therefore, ATF-2 is absolutely required for the full induction of E-selectin and cannot be replaced by other ATF/CREB family members.

The overt defects in the development of bone and the central nervous system in ATF-2 mutant mice could not be predicted from the known distribution and actions of ATF-2, although the diversity of ATF-2 interacting proteins (c-Jun, NF-κB, HMG I(Y), pRb, adenovirus E1A, HTLV-1 Tax and hepatitis B virus

X protein¹³⁻¹⁹) and potential target genes has suggested a widespread role for ATF-2 in gene regulation. Lack of ATF-2 does not substantially affect the expression of other related transcription factor genes, CREB, ATF-1, c-Jun, or c-Fos (Fig. 1h), nor the extent of AP-1 binding (not shown). Indeed, several promoters that bound ATF-2 *in vitro*, including interferon-β, TCR α and β, CD3 δ, and major histocompatibility complex (MHC) class II Aα showed normal mRNA levels in ATF-2 -/- mice (not shown). The bone lesions seen in ATF-2 mutant mice suggest significant similarities between ATF-2 and c-Fos: both are leucine zipper proteins that dimerize with c-Jun, and the deletion of both leads to decreased postnatal survival, runting, and abnormal bone development²⁰⁻²³. In the brains of ATF-2 -/- mice, abnormalities included significant structural defects in the cerebellum, the inner ear and in ventricle size, indicating a widespread and vital role for

FIG. 2 Abnormal skeletal growth in ATF-2 -/- mice. a, X-ray of a severely affected 3-month-old -/- male (bottom) and a +/- male littermate (top). b, *In situ* hybridization of ribs from a day E14.5 +/- embryo using an antisense ATF-2 probe, showing expression of ATF-2 mRNA in resting and proliferating chondrocytes. Darkfield illumination, ×25 original magnification. c, The same field under standard illumination, H and E stain. d, e, Proximal tibial epiphyseal growth plates of 5-month-old +/- (d) and -/- (e) female mice, ×120 original magnification. The upper bracket indicates the physis not invaded by capillaries; the lower bracket delimits the primary cartilaginous trabeculae of the physis. f, ³H-thymidine labelling of growth-plate cartilage. The percentage of labelled nuclei was determined in the upper tibia and lower femur of 3-month-old female +/- and -/- mice. g, Survival to weaning of ATF-2 mutant mice. The offspring from interbreeding of heterozygous mice were counted at birth and genotyped at 3 weeks of age. The expected values were calculated by dividing the total progeny counted according to mendelian ratios.

METHODS. *In situ* hybridization was as described²⁷ using an antisense ATF-2 probe (cDNA positions -327 to -40). Bone specimens were fixed in 10% buffered formalin (0.1M PBS, pH 7.4) for 1-2 weeks, decalcified in 25% formic acid for 3-7 days, embedded in paraffin, cut into 6-µm sections, and stained with safranin O-fast green. *In vivo* DNA labelling of cartilage was by intraperitoneal injection of 20 µCi g⁻¹ body weight ³H-thymidine at 0 and 3 h, followed by collecting at 6 h. Sections (6 µm) were dipped in photographic emulsion, developed at 2 weeks, and counterstained with azure B eosinate. The total number of cells and the labelled cartilage cell nuclei were counted in epiphyseal growth plate serial sections. Statistical significance was determined by Student's t test.



(f) ³ H-thymidine labelling of growth-plate cartilage					
		Labelled nuclei (%)			
		Upper tibia		Lower femur	
+/+		9.69 ± 0.84		23.8 ± 1.1	
-/-		6.11 ± 0.61		14.7 ± 1.0	
Significance		P < 0.001		P < 0.001	

(g) Survival to weaning of ATF-2 mutant mice					
Mating		(+/+)	(+/-)	(-/-)	Dead
(+/-) × (+/-)	Expected	108	215	108	
	Actual	109	163	55	104
	Per cent of expected	100.9	75.8	50.9	

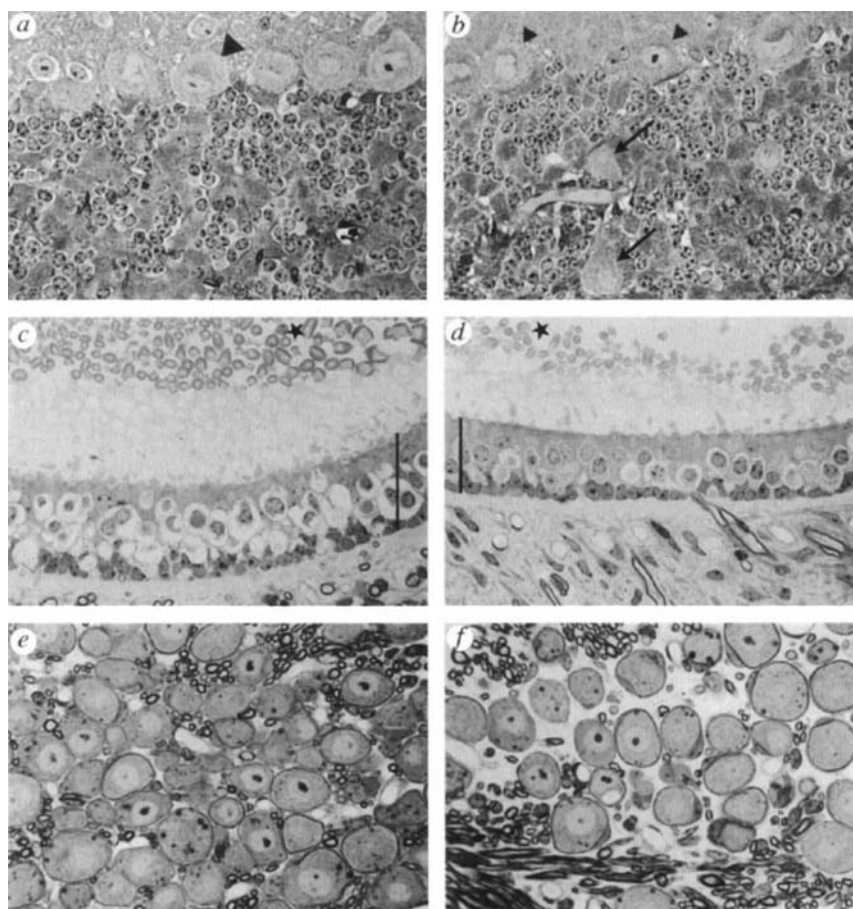


FIG. 3 Central nervous system abnormalities in ATF-2 $-/-$ mice. *a*, Cerebellar cortex from a 2.5-month-old $+/+$ control, showing Purkinje cell bodies (arrowhead) in typical linear position at the molecular-granular cell layer boundary. *b*, The corresponding region from $-/-$ littermate has fewer Purkinje cell bodies (arrowheads). Large neuron cell bodies, similar in size and shape to Purkinje cells, appear in the granular layer (arrows). *c*, Utricular macule of the inner ear from a 9.5-month-old $+/+$ mouse showing the thickness of the sensory epithelium (bar) and the density of the otoconia (asterisk), which are reduced in the $-/-$ littermate (*d*). *e*, *f*, Closely packed neuron cell bodies in the vestibular ganglion of a $+/+$ mouse (*e*) contrast with the looser packing in a $-/-$ littermate (*f*). **METHODS.** Mice were anaesthetized with Avertin (0.02 ml per body weight) and tissues were prepared as described²⁸. Original magnification $\times 400$ (*a*, *b*) and $\times 500$ (*c*–*f*).

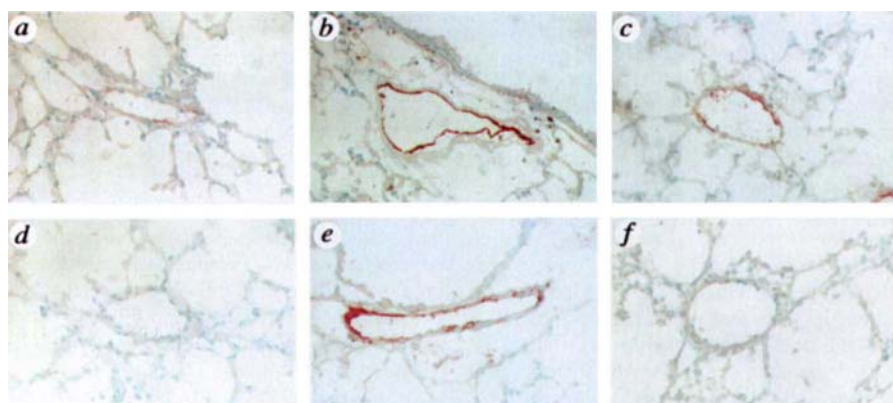


FIG. 4 Induction of E-selectin in mouse lung. *a*, Immunohistochemistry of untreated $-/-$ mouse, anti-VCAM-1 antibody. *b*, LPS-treated $+/+$ mouse, anti-VCAM-1. *c*, LPS-treated $-/-$ mouse, anti-VCAM-1. *d*, Untreated $-/-$ mouse, anti-E-selectin. *e*, LPS-treated $+/+$ mouse, anti-E-selectin. *f*, LPS-treated $-/-$ mouse, anti-E-selectin.

METHODS. Appropriate mice were injected intraperitoneally with 50 μ g LPS and 3 h later lungs were collected and inflated with 50% OCT compound (Miles). Sections of mouse lung were incubated for 1 h as described²⁹ with a 1:40 dilution in PBS of antibody M/K to mouse VCAM-1 or a 1:10 dilution in PBS of antibody 9A9 to mouse E-selectin and stained red when reactive. Original magnification $\times 1,000$.

ATF-2 in central nervous system development. Although ATF-2 is a member of a large gene family with a common recognition sequence, its role is unique and not duplicated, with dramatic, pleiotropic and unexpected effects on development. □

Received 20 July; accepted 10 November 1995.

1. Craig, A. M. & Denhardt, D. T. *Gene* **100**, 163–171 (1991).
2. Matsuura, S., Kishi, F. & Kajiji, T. *Biochem. Biophys. Res. Commun.* **168**, 993–1000 (1990).
3. Campbell, C., Flenniken, A., Skup, D. & Williams, B. J. *biol. Chem.* **266**, 7199–7206 (1991).
4. Schule, R. et al. *Cell* **61**, 497–504 (1990).
5. Le Merrer, M. et al. *Nature Genet.* **6**, 318–321 (1994).
6. Rousseau, F. et al. *Nature* **371**, 252–254 (1994).
7. Shiang, R. et al. *Cell* **78**, 335–342 (1994).
8. Kara, C. J., Liou, H.-C., Ivashkiv, L. B. & Glimcher, L. H. *Molec. cell. Biol.* **10**, 1347–1357 (1990).
9. Takeda, J. et al. *Oncogene* **6**, 1009–1014 (1991).
10. Kaszubska, W. et al. *Molec. cell. Biol.* **13**, 7180–7190 (1993).
11. Schindler, U. & Baichwal, V. R. *Molec. cell. Biol.* **14**, 5820–5831 (1994).
12. Whitley, M. Z., Thanos, D., Read, M. A., Maniatis, T. & Collins, T. *Molec. cell. Biol.* **14**, 6464–6475 (1994).

13. Ivashkiv, L. B. et al. *Molec. cell. Biol.* **10**, 1609–1621 (1990).
14. Du, W., Thanos, D. & Maniatis, T. *Cell* **74**, 887–898 (1993).
15. Hooft van Huijsduijnen, R. et al. *J. biol. Chem.* **267**, 22385–22391 (1992).
16. Kim, S.-J. et al. *Nature* **358**, 331–334 (1992).
17. Liu, F. & Green, M. R. *Cell* **61**, 1217–1224 (1990).
18. Franklin, A. A. et al. *J. biol. Chem.* **268**, 21225–21231 (1993).
19. Maguire, H. F., Hoeftler, J. P. & Siddiqui, A. *Science* **252**, 842–844 (1991).
20. Bourchuladze, R. et al. *Cell* **79**, 59–68 (1994).
21. Hummler, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 5647–5651 (1994).
22. Johnson, R. S., Spiegelman, B. M. & Papaioannou, V. *Cell* **71**, 577–586 (1992).
23. Wang, Z.-Q. et al. *Nature* **360**, 741–745 (1992).
24. Maekawa, T. et al. *EMBO J.* **8**, 2023–2028 (1989).
25. Grusby, M. J., Johnson, R. S., Papaioannou, V. E. & Glimcher, L. H. *Science* **253**, 1417–1420 (1991).
26. Gupta, S., Campbell, D., Derjard, B. & Davis, R. J. *Science* **267**, 389–393 (1995).
27. Clauss, I. M. et al. *Dev. Dynamics* **197**, 146–156 (1993).
28. Hayat, M. A. (ed.) *Principles and Techniques of Electron Microscopy: Biological Applications* 3rd edn (CRC Press, Boca Raton, 1989).
29. Fries, J. W. U. et al. *Am. J. Pathol.* **143**, 725–737 (1993).

ACKNOWLEDGEMENTS. We thank L. Williams and D. Vestweber for gifts of probes, P. Kincade and B. Wolitsky for gifts of antibodies and J. Kim for help with EMSA. This work was supported by grants from the NIH, the H. and L. Mather Foundation, and the Peabody Foundation.

Adhesion through L-selectin requires a threshold hydrodynamic shear

Erik B. Finger*, Kamal D. Puri*, Ronen Alon*, Michael B. Lawrence†, Ulrich H. von Andrian* & Timothy A. Springer*‡

* Center for Blood Research, and Department of Pathology, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

† Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia 22903, USA

SELECTINS are cell adhesion molecules that bind carbohydrate ligands and promote interaction between leukocytes and the vessel wall in vascular shear flow^{1,2}. Selectin–ligand bonds have high mechanical strength, allowing initial tethering to the vessel wall through one or few bonds, and have fast on and off rates, permitting rolling in response to hydrodynamic drag³. The L-selectin molecule on leukocytes binds to peripheral node addressin on high endothelial venules of lymph nodes to mediate leukocyte rolling^{4,5} and binds to a ligand on neutrophils to mediate rolling of leukocytes over one another⁶. Here we describe a surprising mechanism for regulation of these interactions, both *in vitro* and *in vivo*. Shear above a critical threshold is required to promote and maintain rolling interactions through L-selectin, but not through E-selectin, P-selectin or VCAM-1. The shear threshold requirement for L-selectin may be physiologically important in low shear to prevent inappropriate aggregation of leukocytes and interaction with the vessel wall.

When neutrophils or T lymphocytes were perfused through a flow chamber containing purified peripheral node addressin (PNAd) or the CD34 component of PNAd incorporated in a phosphatidylcholine bilayer on one wall, maximal accumulation of rolling cells was found at a wall shear stress of 1 dyn cm^{-2} (Fig. 1a). Accumulation dropped above 2 dyn cm^{-2} . However, there was also an unexpected decrease in accumulation of rolling cells below 0.6 dyn cm^{-2} . By contrast, efficient accumulation of rolling neutrophils was seen on E-selectin and P-selectin at the lowest shear stresses tested (Fig. 1a).

Shear above a threshold value was also required for formation of stable tethers, defined as newly initiated tethers to the PNAd substrate that persist as rolling adhesions for $> 0.5 \text{ s}$ (Fig. 1b). The number of stable tethers was divided by the number of cells that flowed through the field of view to correct for lower flux at lower shear; nonetheless, a threshold of about 0.4 dyn cm^{-2} was seen for stable tethering to PNAd and CD34 at all site densities tested (Fig. 1b and data not shown). By contrast, stable tethers on E-selectin, P-selectin, and VCAM-1 were most frequent at low shear (Fig. 1c). No stable tethers formed when leukocytes settled in stasis on PNAd substrates, because when flow was started at 1.8 dyn cm^{-2} , the cells travelled a few cell diameters at the hydrodynamic velocity (the velocity of a non-interacting cell near the wall) before they tethered and slowed in motion as they began rolling (our unpublished results and E. Butcher, personal communication).

The shear dependence of rolling adhesions through L-selectin was dramatically illustrated by counterintuitive shifts in leukocyte velocity when the shear was shifted between values below and above the threshold. The hydrodynamic velocity of leukocytes was measured in EDTA, which inhibits Ca^{2+} -dependent selectin interactions⁸. The velocity in EDTA and Ca^{2+} was similar on PNAd below 0.4 dyn cm^{-2} (Fig. 2a). However, when shear stress

was increased to 0.73 dyn cm^{-2} , most of the cells near the wall tethered and rolled at lower velocity (Fig. 2a). When lymphocytes rolling at $16.2 \pm 1.0 \mu\text{m s}^{-1}$ on PNAd at 1.8 dyn cm^{-2} were subjected to a rapid reduction in shear stress to 0.18 dyn cm^{-2} , they detached from the PNAd substrate and moved at $41.3 \pm 2.3 \mu\text{m s}^{-1}$ (the increase in velocity was significant, with $P < 0.001$ (Fig. 2b). The velocity at 0.18 dyn cm^{-2} matched that of lymphocytes in EDTA, showing no interaction with the substrate. When flow was resumed at a wall shear stress of 1.8 dyn cm^{-2} , the lymphocytes retethered and resumed rolling at lower velocity. Similar experiments with neutrophils are illustrated by superimposition of 10 images captured at 0.1-s intervals (Fig. 2c–f). In Ca^{2+} , neutrophils are rollingly adherent at 1.5 dyn cm^{-2} (Fig. 2c), but after shift to 0.22 dyn cm^{-2} the cells detach and move faster (Fig. 2d). The hydrodynamic velocities of neutrophils in EDTA at 1.5 dyn cm^{-2} (Fig. 2e; only 3 to 5 images of each cell are within the field of view) are much faster than in Ca^{2+} (Fig. 2c), but at 0.22 dyn the velocities of cells in EDTA (Fig. 2f) and in Ca^{2+} (Fig. 2d) are comparable. Reversible rolling at 1.5 dyn cm^{-2} and detachment at 0.22 dyn cm^{-2} was seen for at least 20 cycles of shifts in flow velocity, and behaviour of the cells at the new shear stress equilibrated in less than 1 s.

By contrast, when lymphocytes or neutrophils rolling on E-selectin or P-selectin, or VLA-4 transfectants rolling on VCAM-1 at $\geq 1.5 \text{ dyn cm}^{-2}$ were subjected to a decrease in shear to 0.15 or 0.3 dyn cm^{-2} , they rolled more slowly, and none of the cells detached (Fig. 3). Similarly, L-selectin-dependent rolling on sialyl Lewis^x and fucoidan substrates did not exhibit a threshold; rolling slowed with no detachment when shear was reduced (Fig. 3). We suspect that this reflects the very high density of these ligands in the assay, that is, $12,000 \text{ sites } \mu\text{m}^{-2}$ for sialyl Lewis^x; the density required to support rolling is about 100-fold higher for sialyl Lewis^x than for the CD34 component of PNAd^{9,10}.

As previously reported for neutrophils adherent to endothelial cells⁶, we found that a layer of neutrophils adherent to and rolling slowly on E-selectin provided a substrate for tethering and rolling of further neutrophils. These neutrophil–neutrophil interactions were L-selectin-dependent (ref. 6 and data not shown). Furthermore, they were shear-dependent. Neutrophils rolled jerkily on other neutrophils at 1.5 dyn cm^{-2} , with a velocity of $127 \pm 24 \mu\text{m s}^{-1}$ (Fig. 2g); however, when shear was reduced to 0.3 dyn cm^{-2} the cells detached and moved smoothly at increased velocity of $197 \pm 44 \mu\text{m s}^{-1}$ ($P < 0.001$) (Figs 2h, 3).

Rolling in murine Peyer's patch HEV¹¹ was tested to determine whether a threshold shear was required for L-selectin-dependent interactions *in vivo*. Human neutrophils treated with neuraminidase to remove E- or P-selectin-dependent adhesion formed rolling adhesions in Peyer's patch HEV at physiological shear (Fig. 4). These interactions were $97 \pm 6.9\%$ inhibitable by L-selectin monoclonal antibody (not shown). To assess rolling at low shear, the vascular supply to the Peyer's patch was partially restricted and neutrophil behaviour in the same vessels was observed. The rolling fraction in the hypoperfusion state was decreased in 14/14 vessels examined, by a mean $\pm \text{s.d.}$ of $89 \pm 19\%$. The rolling fraction increased back to baseline levels when normal flow was reinstated (not shown). Plotting all the data (Fig. 4) revealed a bell-shaped dependence on venular shear rate with a threshold similar to that observed *in vitro*. Rolling *in vivo* compared to *in vitro* is favoured by more efficient leukocyte margination owing to contributions by red blood cells and other microhaemodynamic factors¹². Thus, rolling was observed within a range of shear stresses from ~ 2.5 to 40 dyn cm^{-2} , assuming a blood viscosity of 0.025 poise.

Previously, cell adhesion has been reported to be upregulated by increased density on the cell surface, affinity, or clustering of adhesion molecules¹³. Increased shear is an unexpected and unprecedented mechanism for upregulating cell adhesion, and so far appears to be unique to L-selectin. A dramatic and counterintuitive illustration of the phenomenon we have described is that as the flow rate is increased to yield a wall shear stress above the

‡ To whom correspondence should be addressed.

threshold, the velocity of leukocytes slows as they tether and roll. Active regulation by cellular signalling appears not to be required, as the phenomenon occurs within less than 1 s, and is repeated through many cycles of shifts in shear. Shear above a threshold is required for both initiation and maintenance of rolling. Loss of rolling interactions within less than 1 s after a drop in shear below the threshold implies a lifetime for the L-selectin-carbohydrate bond of less than 1 s; indeed, recent measurements show a lifetime of 0.15 s (R.A. *et al.*, manuscript in preparation). 'Catch bonds' are theoretically possible that strengthen when tensile force is applied¹⁴; however, this is not the explanation, because tensile force shortens the lifetime of the L-selectin bond (R.A. *et al.*, in preparation). Transient tethers with a lifetime of 0.15 s occur below the shear threshold (R.A. *et al.*, in preparation) but do not progress to rolling adhesions or the stable tethers of duration > 0.5 s measured here. We suspect that once the first short-lived L-selectin bond forms between the cell and the substrate, hydrodynamic shear may provide what is known in engineering as

transport, to promote the formation of additional L-selectin bonds with the substrate before the first bond dissociates. L-selectin is clustered on the tips of microvilli¹⁵⁻¹⁷ and shear-driven transport of the cell may bring further microvilli in contact with the substrate. The short lifetime of L-selectin (R.A. *et al.*, in preparation) compared with P-selectin³ and E-selectin¹⁸ bonds is consistent with the uniqueness of the shear requirement to L-selectin and the postulated importance of short bond lifetime in the shear threshold phenomenon.

Regulation by hydrodynamic shear of adhesion through L-selectin is likely to be biologically important. Interactions through all other known adhesion molecules are enhanced by low shear; thus, loss of L-selectin interactions may provide a counterbalance to prevent inappropriate interactions with the vessel wall in low shear. This may be important in vessels with inherently low shear, such as sinusoids and large veins; in pathologies with hypoperfusion; and in the ascending aorta, which is prone to atherosclerosis and where flow reverses and shear goes briefly to

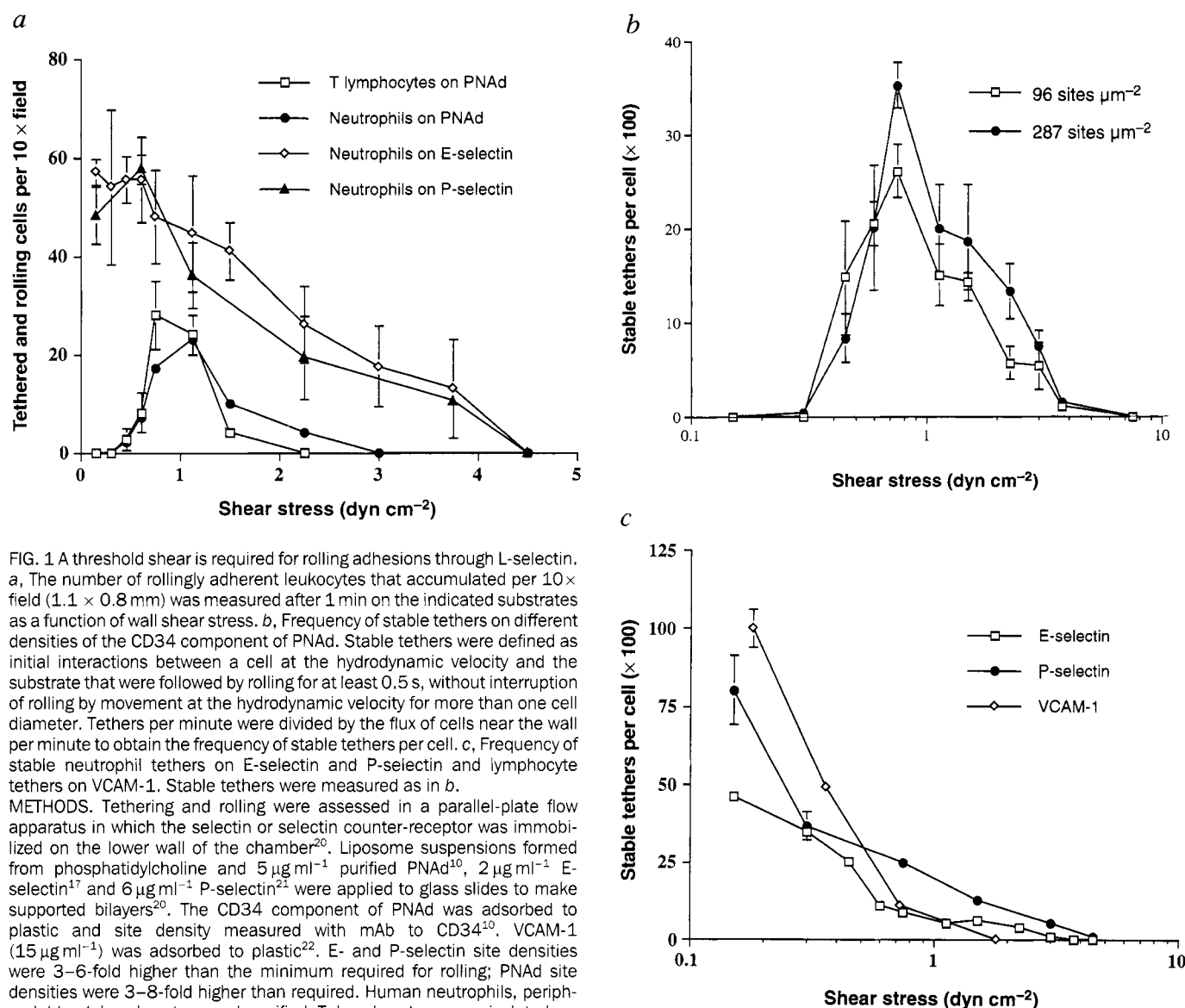


FIG. 1 **a** A threshold shear is required for rolling adhesions through L-selectin. **a**, The number of rollingly adherent leukocytes that accumulated per $10 \times$ field ($1.1 \times 0.8 \text{ mm}$) was measured after 1 min on the indicated substrates as a function of wall shear stress. **b**, Frequency of stable tethers on different densities of the CD34 component of PNAd. Stable tethers were defined as initial interactions between a cell at the hydrodynamic velocity and the substrate that were followed by rolling for at least 0.5 s, without interruption of rolling by movement at the hydrodynamic velocity for more than one cell diameter. Tethers per minute were divided by the flux of cells near the wall per minute to obtain the frequency of stable tethers per cell. **c**, Frequency of stable neutrophil tethers on E-selectin and P-selectin and lymphocyte tethers on VCAM-1. Stable tethers were measured as in **b**.

METHODS. Tethering and rolling were assessed in a parallel-plate flow apparatus in which the selectin or selectin counter-receptor was immobilized on the lower wall of the chamber²⁰. Liposome suspensions formed from phosphatidylcholine and $5 \mu\text{g ml}^{-1}$ purified PNAd¹⁰, $2 \mu\text{g ml}^{-1}$ E-selectin¹⁷ and $6 \mu\text{g ml}^{-1}$ P-selectin²² were applied to glass slides to make supported bilayers²⁰. The CD34 component of PNAd was adsorbed to plastic and site density measured with mAb to CD34¹⁰. VCAM-1 ($15 \mu\text{g ml}^{-1}$) was adsorbed to plastic²². E- and P-selectin site densities were 3–6-fold higher than the minimum required for rolling; PNAd site densities were 3–8-fold higher than required. Human neutrophils, peripheral blood lymphocytes, and purified T lymphocytes were isolated as described^{23,24} and perfused through the flow chamber at 5×10^5 cells per ml in Ca^{2+} - and Mg^{2+} -free Hank's balanced salt solution supplemented with 10 mM HEPES, pH 7.4, 2 mM Ca^{2+} , 0.25% human serum albumin. Experiments were videotaped for later analysis. Specificity of interaction was confirmed and the hydrodynamic velocity was determined by addition of

5 mM EDTA; selectin interactions are Ca^{2+} dependent⁸. Cells accumulating on or forming stable tethers to E-selectin or P-selectin or forming stable tethers to PNAd after first interacting with a previously adherent cell were not counted. Error bars show s.d., $n = 3-6$.

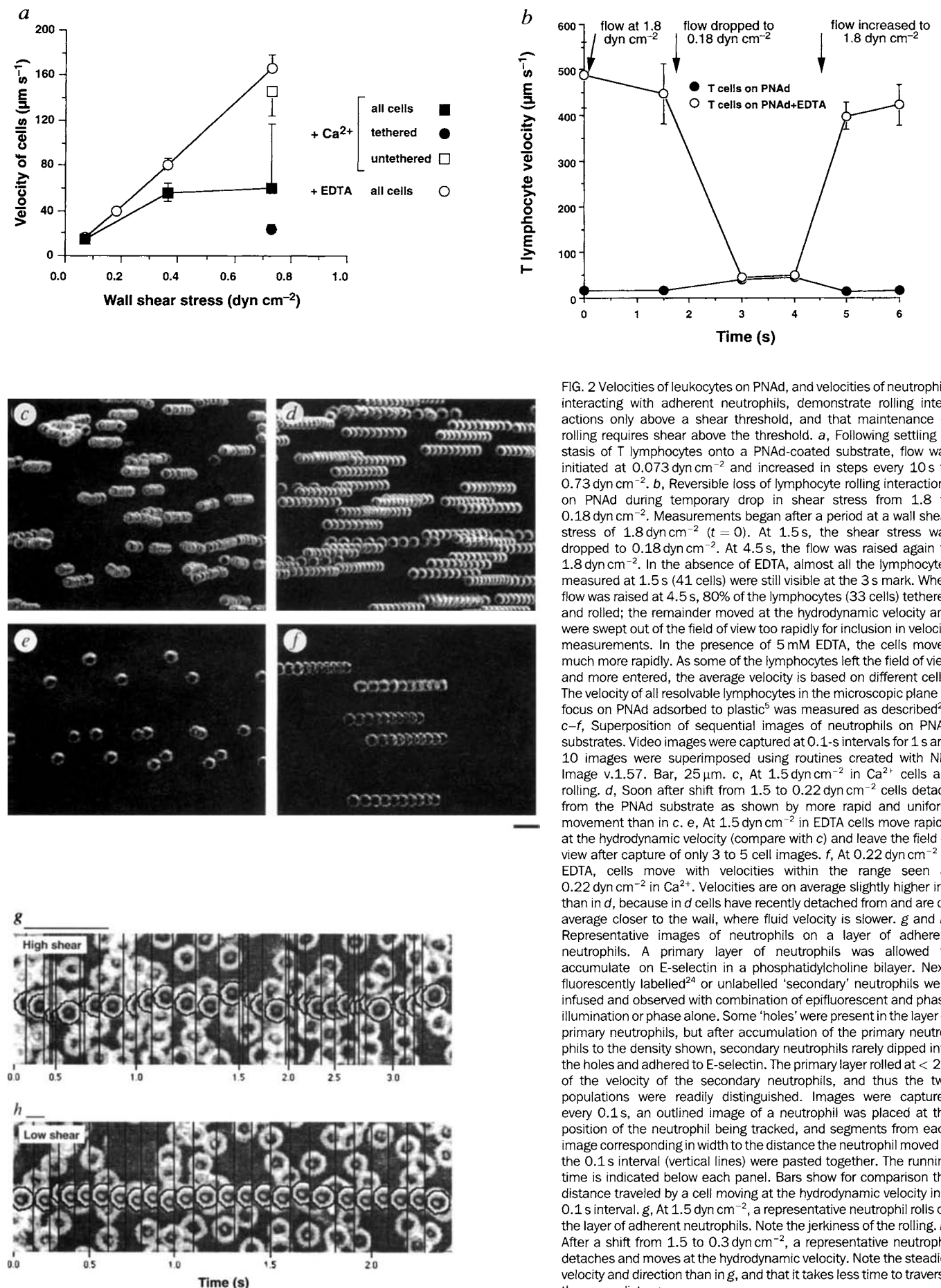


FIG. 2 Velocities of leukocytes on PNAd, and velocities of neutrophils interacting with adherent neutrophils, demonstrate rolling interactions only above a shear threshold, and that maintenance of rolling requires shear above the threshold. **a**, Following settling in stasis of T lymphocytes onto a PNAd-coated substrate, flow was initiated at $0.073 \text{ dyn cm}^{-2}$ and increased in steps every 10 s to 0.73 dyn cm^{-2} . **b**, Reversible loss of lymphocyte rolling interactions on PNAd during temporary drop in shear stress from 1.8 to 0.18 dyn cm^{-2} . Measurements began after a period at a wall shear stress of 1.8 dyn cm^{-2} ($t = 0$). At 1.5 s , the shear stress was dropped to 0.18 dyn cm^{-2} . At 4.5 s , the flow was raised again to 1.8 dyn cm^{-2} . In the absence of EDTA, almost all the lymphocytes measured at 1.5 s (41 cells) were still visible at the 3 s mark. When flow was raised at 4.5 s , 80% of the lymphocytes (33 cells) tethered and rolled; the remainder moved at the hydrodynamic velocity and were swept out of the field of view too rapidly for inclusion in velocity measurements. In the presence of 5 mM EDTA, the cells moved much more rapidly. As some of the lymphocytes left the field of view and more entered, the average velocity is based on different cells. The velocity of all resolvable lymphocytes in the microscopic plane of focus on PNAd adsorbed to plastic⁵ was measured as described²⁰. **c–f**, Superposition of sequential images of neutrophils on PNAd substrates. Video images were captured at 0.1-s intervals for 1 s and 10 images were superimposed using routines created with NIH Image v.1.57. Bar, $25 \mu\text{m}$. **c**, At 1.5 dyn cm^{-2} in Ca^{2+} cells are rolling. **d**, Soon after shift from 1.5 to 0.22 dyn cm^{-2} cells detach from the PNAd substrate as shown by more rapid and uniform movement than in **c**. **e**, At 1.5 dyn cm^{-2} in EDTA cells move rapidly at the hydrodynamic velocity (compare with **c**) and leave the field of view after capture of only 3 to 5 cell images. **f**, At 0.22 dyn cm^{-2} in EDTA, cells move with velocities within the range seen at 0.22 dyn cm^{-2} in Ca^{2+} . Velocities are on average slightly higher in **f** than in **d**, because in **d** cells have recently detached from and are on average closer to the wall, where fluid velocity is slower. **g** and **h**, Representative images of neutrophils on a layer of adherent neutrophils. A primary layer of neutrophils was allowed to accumulate on E-selectin in a phosphatidylcholine bilayer. Next, fluorescently labelled²⁴ or unlabelled 'secondary' neutrophils were infused and observed with combination of epifluorescent and phase illumination or phase alone. Some 'holes' were present in the layer of primary neutrophils, but after accumulation of the primary neutrophils to the density shown, secondary neutrophils rarely dipped into the holes and adhered to E-selectin. The primary layer rolled at $< 2\%$ of the velocity of the secondary neutrophils, and thus the two populations were readily distinguished. Images were captured every 0.1 s , an outlined image of a neutrophil was placed at the position of the neutrophil being tracked, and segments from each image corresponding in width to the distance the neutrophil moved in the 0.1 s interval (vertical lines) were pasted together. The running time is indicated below each panel. Bars show for comparison the distance traveled by a cell moving at the hydrodynamic velocity in a 0.1 s interval. **g**, At 1.5 dyn cm^{-2} , a representative neutrophil rolls on the layer of adherent neutrophils. Note the jerkiness of the rolling. **h**, After a shift from 1.5 to 0.3 dyn cm^{-2} , a representative neutrophil detaches and moves at the hydrodynamic velocity. Note the steadier velocity and direction than in **g**, and that it takes less time to traverse the same distance.

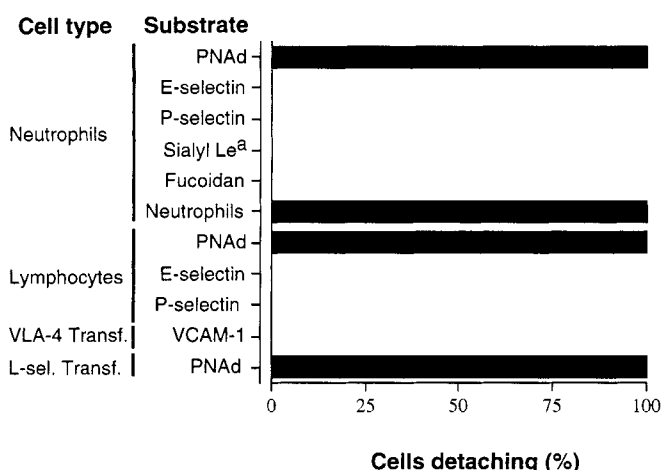
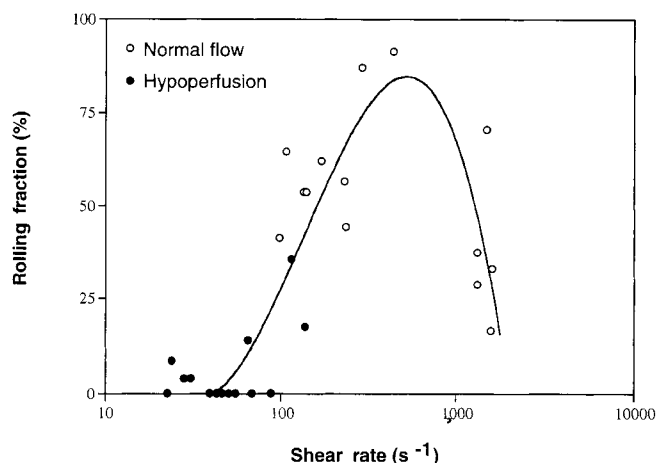


FIG. 3 Lymphocytes and neutrophils require hydrodynamic shear for maintenance of rolling on physiological L-selectin ligands, but not for other selectin-mediated rolling interactions. Cells rolling on the indicated substrates at $> 1.5 \text{ dyn cm}^{-2}$ were subjected to reduction in shear to 0.15 or 0.3 dyn cm^{-2} and the percentage of cells was determined that detached after 10 s of reduced flow, as shown by movement at the hydrodynamic velocity. (2 to 6 trials for each experimental group; at least 50 cells scored per group).

METHODS. Lymphocytes, neutrophils, K562 cell transfectants expressing VLA-4 (ref. 25), or K562 transfectants expressing L-selectin (unpublished) were allowed to tether to the indicated substrates in stasis or at 0.75 dyn cm^{-2} , flow was increased to 1.5 to 35 dyn cm^{-2} , and then decreased to 0.15 or 0.3 dyn cm^{-2} to measure detachment. Sialyl Lewis^a glycolipid ($1 \mu\text{g ml}^{-1}$)⁹, fucoidan-BSA ($100 \mu\text{g ml}^{-1}$)²⁶, VCAM-1 ($15 \mu\text{g ml}^{-1}$)²², and PNAAd ($1.5 \mu\text{g ml}^{-1}$)¹⁰ were adsorbed at the indicated concentration on plastic. E-selectin ($2 \mu\text{g ml}^{-1}$) and P-selectin ($6 \mu\text{g ml}^{-1}$) were prepared in bilayers as described above. Rolling of neutrophils on adherent neutrophils was as in Fig. 2.

FIG. 4 A threshold hydrodynamic shear is required for neutrophil rolling *in vivo*. Rolling of neuraminidase-treated human neutrophils on mouse Peyer's patch HEV was observed intravital¹¹. The fraction of rolling cells in the normal flow state (open circles) was compared with the rolling fraction in the same vessels after hypoperfusion was achieved (filled circles).

METHODS. Mouse small bowel was pulled out through a midline abdominal incision, placed in a ring of vacuum grease built onto a microscope slide, wetted with saline, and gently covered with a coverslip. Peyer's patches were identified and observed by intravital microscopy. Isolated human neutrophils were treated with 0.1 U ml^{-1} neuraminidase as described¹⁷ and labelled with 2',7'-bis(2-carboxyethyl)-5-(and-6-)carboxyfluorescein, acetoxymethyl ester. These cells were resuspended in L15 medium at $5-10 \times 10^7 \text{ ml}^{-1}$, treated with CD18 mAb IB4 (ref. 27) and injected via a carotid cannula. Neutrophil interactions were observed with stroboscopic epi-illumination using a low-lag SIT camera (Dage MTI) and recorded on videotape. Hypoperfusion was achieved by transiently clamping of branches of the mesenteric artery directly supplying the Peyer's patch and by partial ligation of the bowel loop oral and aboral to the area of observation to reduce collateral flow. Specificity of interaction was assessed with an aliquot of neutrophils pretreated with L-selectin mAb Dreg 56 (ref. 28). Diameters of observed vessels were determined after i.v. injection of $250 \mu\text{l}$ of a 10 mg ml^{-1} FITC-dextran (M_r 150,000) solution. Rolling cells were identified by a qualitative assessment of their interaction with the vessel wall; rolling cells show multiple discrete slowings while in contact with the endothelial layers, whereas free cells transit the vessel smoothly without obvious interaction. The rolling fraction was determined as described²⁹. Free cell velocities were measured with a computer-aided image analysis system and used to estimate shear rate²⁹. Data represent measurements from three separate experiments.



zero during every heartbeat. Neutrophils express both L-selectin and L-selectin ligands^{6,19} and would tend to aggregate through L-selectin unless this interaction was regulated by shear. Notably,

shear in blood vessels is highest at the wall and approaches zero towards the centre of the vessel, restricting neutrophil-neutrophil interactions to regions near the wall. □

Received 20 July; accepted 8 November 1995.

- Springer, T. A. *Rev. Physiol.* **57**, 827-872 (1995).
- Springer, T. A. *Traffic Signals and Rolling Cells, a Videotape for Teaching* (Springer, Boston, 1995).
- Alon, R., Hammer, D. A. & Springer, T. A. *Nature* **374**, 539-542 (1995).
- Berg, E. L., Robinson, M. K., Warnock, R. A. & Butcher, E. C. *J. Cell Biol.* **114**, 343-349 (1991).
- Lawrence, M. L., Berg, E. L., Butcher, E. C. & Springer, T. A. *Eur. J. Immun.* **25**, 1025-1031 (1995).
- Bargatze, R. F., Kurk, S., Butcher, E. C. & Jutila, M. A. *J. exp. Med.* **180**, 1785-1792 (1994).
- Goldman, A. J., Cox, R. G. & Brenner, H. *Chem. Engng Sci.* **22**, 653-660 (1967).
- Lasky, L. A. *Science* **258**, 964-969 (1992).
- Alon, R., Feizi, T., Yuen, C.-T., Fuhlbrigge, R. C. & Springer, T. A. *J. Immun.* **154**, 5356-5366 (1995).
- Puri, K. D., Finger, E. B., Gaudernack, G. & Springer, T. A. *J. Cell Biol.* **131**, 261-270 (1995).
- Bargatze, R. F., Jutila, M. A. & Butcher, E. C. *Immunology* **3**, 99-108 (1995).
- Schmid-Schönbein, G. W., Usami, S., Skalak, R. & Chien, S. *Microvasc. Res.* **19**, 45-70 (1980).
- Springer, T. A. *Nature* **346**, 425-433 (1990).
- Dembo, M., Tórnay, D. C., Saxman, K. & Hammer, D. *Proc. R. Soc. Lond. B* **234**, 55-83 (1988).
- Pickar, L. J. *et al. Cell* **66**, 921-933 (1991).

- Erlandsen, S. L., Hasslen, S. R. & Nelson, R. D. *J. Histochem. Cytochem.* **41**, 327-333 (1993).
- Lawrence, M. B., Bainton, D. F. & Springer, T. A. *Immunity* **1**, 137-145 (1994).
- Kaplanski, G. *et al. Biophys. J.* **64**, 1922-1933 (1993).
- Rochon, Y. P., Simon, S. I., Lynam, E. B. & Sklar, L. A. *J. Immun.* **152**, 1385-1393 (1994).
- Lawrence, M. B. & Springer, T. A. *Cell* **65**, 859-873 (1991).
- Moore, K. L., Varki, A. & McEver, R. P. *J. Cell Biol.* **112**, 491-499 (1991).
- Pepinsky, B. *et al. J. Biol. Chem.* **267**, 17820-17826 (1992).
- Miller, L. J., Bainton, D. F., Borregaard, N. & Springer, T. A. *clin. Invest.* **80**, 535-544 (1987).
- Carr, M. W., Roth, S. J., Luther, E., Rose, S. S. & Springer, T. A. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3652-3656 (1994).
- Alon, R. *et al. J. Cell Biol.* **128**, 1243-1253 (1995).
- Glabe, C. G., Harty, P. K. & Rosen, S. D. *Analyt. Biochem.* **130**, 287-294 (1983).
- Wright, S. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 5699-5703 (1983).
- Kishimoto, T. K., Jutila, M. A. & Butcher, E. C. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2244-2248 (1990).
- von Andrian, U. H. *et al. Am. J. Physiol.* **263**, H1034-H1044 (1992).

ACKNOWLEDGEMENTS. This work was supported by a grant from the NIH. We thank E. Butcher, K. Kishimoto, R. Lobb and R. McEver for antibodies and proteins.

A bipolar kinesin

A. S. Kashina, R. J. Baskin, D. G. Cole,
K. P. Wedaman, W. M. Saxton* & J. M. Scholey

Section of Molecular and Cellular Biology, University of California, Davis, California 95616, USA

* Department of Biology, Indiana University, Bloomington, Indiana 47405, USA

CHROMOSOME segregation during mitosis depends on the action of the mitotic spindle, a self-organizing, bipolar protein machine which uses microtubules (MTs) and their associated motors^{1,2}. Members of the BimC subfamily of kinesin-related MT-motor proteins are believed to be essential for the formation and functioning of a normal bipolar spindle³⁻¹⁴. Here we report that KRP₁₃₀, a homotetrameric BimC-related kinesin purified from *Drosophila melanogaster* embryos¹³, has an unusual ultrastructure. It consists of four kinesin-related polypeptides assembled into a bipolar aggregate with motor domains at opposite ends, analogous to a miniature myosin filament¹⁵. Such a bipolar 'minifilament' could crosslink spindle MTs and slide them relative to one another. We do not know of any other MT motors that have a bipolar structure.

We investigated the structure of purified embryonic KRP₁₃₀ holoenzymes from *Drosophila melanogaster* (Fig. 1a) using electron microscopy¹⁶⁻¹⁹ (Fig. 2). Single rotary-shadowed KRP₁₃₀ molecules are visible as elongated bipolar structures with an overall length of 95.6 ± 9.8 nm ($n = 131$), consisting of two globular domains connected by a central rod with a length of 61.3 ± 8.3 nm ($n = 130$). The rod domain of KRP₁₃₀ appeared less flexible than that of kinesin, although a small fraction of KRP₁₃₀ molecules were slightly bent in the middle of the rod (average bend angle was 134° , compared with 180° for straight molecules; range 127° – 154°). The dimensions of the globular domains at opposite ends of KRP₁₃₀ molecules were indistinguishable, being 21.7 ± 3.7 nm in diameter ($n = 154$), approximately twice the diameter of a single rotary-shadowed kinesin motor domain^{17,18}. This is consistent with there being two close-packed motor domains at each end of the tetrameric KRP₁₃₀ molecule.

These rotary-shadowed images suggest that KRP₁₃₀ holoenzymes have the structure shown in Fig. 3a. In this model, two KRP₁₃₀ polypeptides associate by coiled-coil interactions to form a parallel dimer. Two dimers then associate by lateral interactions between their coiled-coil regions to form a bipolar tetramer, which consists of a central rod with two motor domains projecting from either end. To test the hypothesis that the bipolar appearance of KRP₁₃₀ molecules reflects the presence of globular MT-binding motor domains at both ends, we first examined MT-KRP₁₃₀ complexes by negative stain and rotary-shadow electron microscopy (Fig. 3b,c). Bipolar, negatively stained KRP₁₃₀ molecules were observed to bind to microtubules by one or both ends (Fig. 3b). Rotary shadowing of MT-KRP₁₃₀ complexes revealed aperiodic MT-MT crossbridges of length 52.7 ± 8.9 nm ($n = 47$), which were absent in preparations of MTs alone (Fig. 3c). However, kinesin itself, which is thought to crosslink vesicles to MTs *in vivo*, forms similar MT-MT crossbridges under some experimental conditions²⁰. It is therefore difficult to interpret our observation that KRP₁₃₀ can form crossbridges to adjacent MTs.

Therefore, to test whether KRP₁₃₀ is a bipolar tetramer, we used electron microscopy to study rotary-shadowed KRP₁₃₀ molecules decorated with an antibody that reacts specifically with the motor domains of Eg5 and its close relatives, all members of the BimC subfamily of kinesins (Fig. 1b,c). We had previously observed that the motor-domain monoclonal IgG, SUK4, decorates only one end of the conventional kinesin molecule¹⁷. In contrast, the Eg5 motor-domain antibodies used here usually bound to both ends of KRP₁₃₀ molecules (60–70% of those molecules examined), decorating the globular domains and resulting in a marked increase in

their dimensions (Fig. 4a). For example, we examined 70 randomly selected KRP₁₃₀ molecules that had been incubated with motor-domain antibody before rotary shadowing, of which 18 were undecorated (overall length, 102 ± 9.9 nm; rod length, 62.6 ± 11.3 nm; globular domain diameter, 22.6 ± 3.7 nm), 6 were decorated on one end only, and 46 were decorated on both ends (overall length, 127 ± 18.7 nm; rod length, 43.1 ± 13.3 nm; decorated globular domain diameter, 44.7 ± 7.2 nm). By compar-

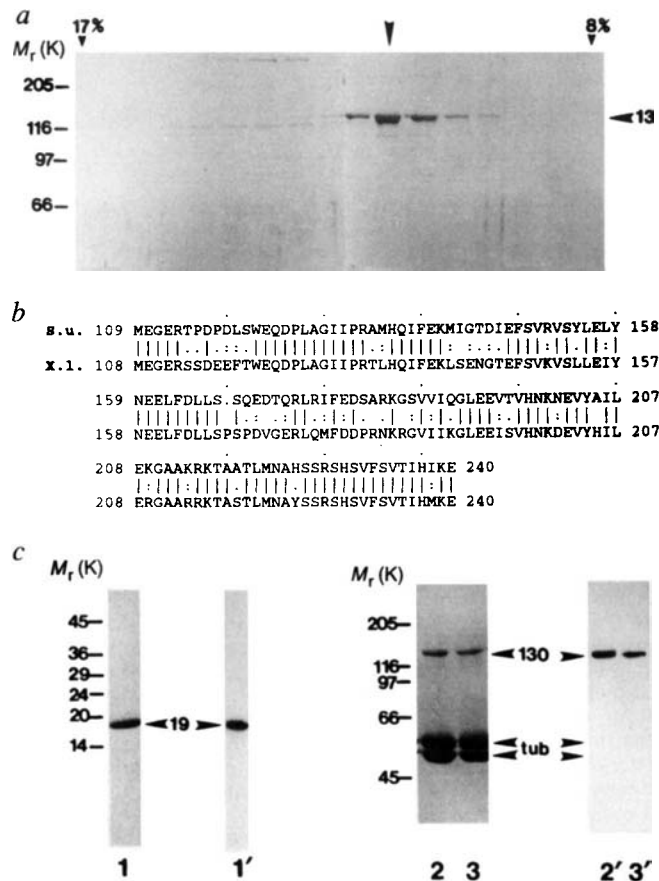


FIG. 1 Purification of KRP₁₃₀ and corresponding motor-domain antibodies for electron microscopy. **a**, SDS-polyacrylamide gel of fractions from sucrose density gradient centrifugation of KRP₁₃₀ (ref. 13). Arrowhead indicates peak fraction. Percentage sucrose (w/v) is indicated for two fractions. **b**, Best-fit comparison of the sea urchin (S.u.) Eg5-related motor-domain subfragment used for raising Eg5 motor-domain antibodies, and the corresponding region of *Xenopus laevis* (X.l.) Eg5 (identity 69%, similarity 83%). **c**, Characterization of the Eg5 motor-domain antibody. SDS-polyacrylamide gels (1–3) and corresponding anti-Eg5 motor-domain immunoblots (1'–3') of a purified fusion protein of *M*_r 19K, (His)₆-SU Eg5 motor, comprising the sea-urchin Eg5-related motor-domain fragment fused to a polyhistidine tag (1,1'), and KRP₁₃₀ pelleted with MTs in the presence of the nonhydrolysable ATP analogue, 5'-adenylyl- β , γ -imidodiphosphate (AMP-PNP) (2,2') or ATP (3,3') in low salt. Numbers on the left refer to *M*_r standards. Arrowheads indicate recombinant motor-domain fragment (19), tubulin (tub) and KRP₁₃₀ (130). **METHODS** KRP₁₃₀ was purified from *Drosophila melanogaster* embryos¹³ and rotary-shadowed directly or following co-pelleting with MTs in AMP-PNP or ATP (KRP₁₃₀ binding to MTs is ATP-sensitive at high but not low ionic strength; Fig. 1c and ref. 13). Antibody specific for KRP₁₃₀ motor domains (Figs 1c and 4a) was raised against a recombinant motor-domain fragment of a sea-urchin egg protein, KRP₁₇₀, that is a close relative of frog Eg5. Expression of a KRP₁₇₀ motor-domain fragment complementary DNA in pRSETB vector yielded a (His)₆-SU Eg5 motor fusion protein of *M*_r 19K, which was purified by Ni-nitrilotriacetic acid metal-chelate affinity chromatography followed by SDS-PAGE electroelution (Fig. 1c, lane 1) and used to raise rabbit polyclonal antibody. Specific antibody was affinity purified from serum using the 19K fusion protein coupled to agarose, analysed by immunoblotting (Fig. 1c) and used for electron microscopy (Fig. 4a).

ison, under similar conditions, Hirokawa and co-workers¹⁸ found that approximately 50% of conventional kinesin molecules were labelled with antibody, and the decorated ends of these molecules displayed an overall diameter of approximately 30 nm, two-thirds the size of the corresponding dimension observed here.

Our results are consistent with the hypothesis that at least two Eg5 motor-domain antibody molecules have decorated each end of a KRP₁₃₀ molecule, and that each KRP₁₃₀ holoenzyme has two motor domains at either end. Thus, based on electron microscopy and previous hydrodynamic studies¹³, we conclude that KRP₁₃₀ holoenzymes are extended, elongated molecules consisting of four motor polypeptides assembled into bipolar 'minifilaments' (Fig. 3a). These 'minifilaments' could crosslink adjacent MTs and slide them relative to one another (Fig. 4b). This model is consistent with recent evidence that an intact carboxy-terminal tail is a prerequisite for Eg5 localization and function in the spindle²¹, presumably because the tail region participates in the self-assembly of functional bipolar holoenzymes.

How could such an assembly function in the spindle? The only other motor protein known to assemble into bipolar aggregates is myosin II (ref. 15), which crosslinks actin filaments and slides

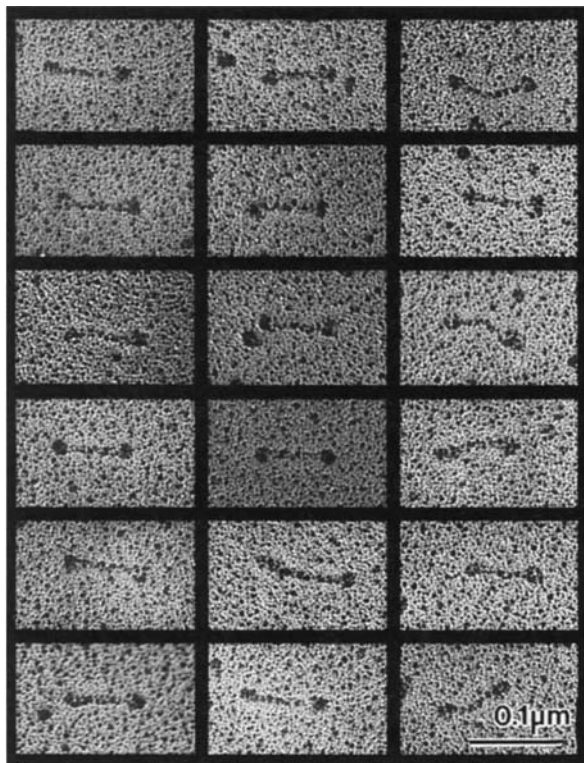


FIG. 2 Electron micrographs of rotary-shadowed KRP₁₃₀ molecules. These images, together with previous studies¹³, are consistent with the hypothesis that a KRP₁₃₀ holoenzyme comprises four 130K subunits assembled into a structure 96 nm long, with a total *M_r* of roughly 500K, Stokes radius of 16.2 nm, and S-value of 7.6S. For comparison, kinesin comprises two heavy chains of 110K–130K and two light chains of 55K–80K assembled into a structure 75–80 nm long with a total *M_r* of 350K–400K, Stokes radius of 9 nm, and S-value of 9.0–9.6S. The dimensions of both rotary-shadowed KRP₁₃₀ and kinesin appear slightly larger than their true dimensions owing to the coating of platinum, which in our experiments is routinely estimated to be approximately 2.5 nm thick.

METHODS. KRP₁₃₀ alone or KRP₁₃₀–MT complexes were mixed with an equal volume of 80% glycerol containing 0.6–1.0 M ammonium acetate and sprayed onto freshly cleaved mica plates. The plates were then vacuum dried, rotary shadowed with platinum at a 6° angle using a Balzers BAF 400T freeze-fracture device, and processed as described previously²⁸. The replicas were visualized on a Philips 410LS electron microscope at 80 kV. No difference in the structure of KRP₁₃₀ was observed with or without the MT binding step.

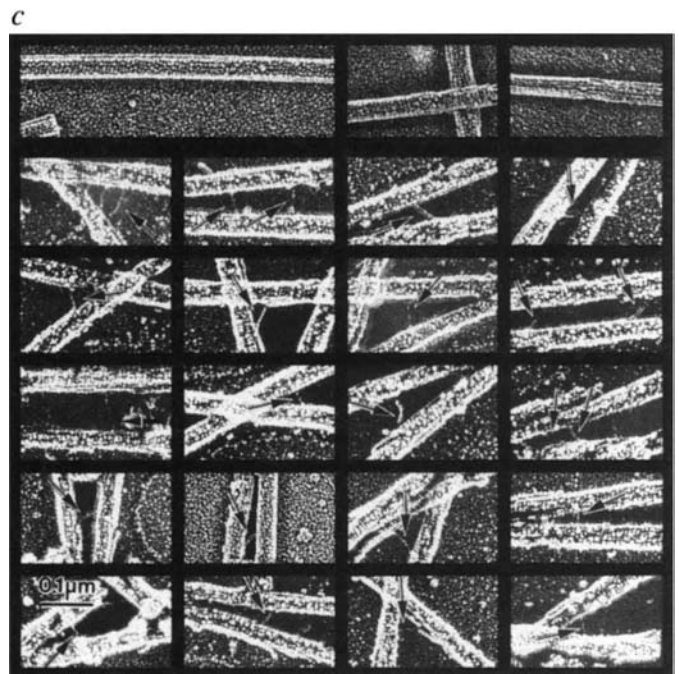
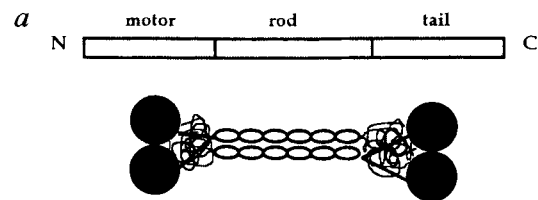
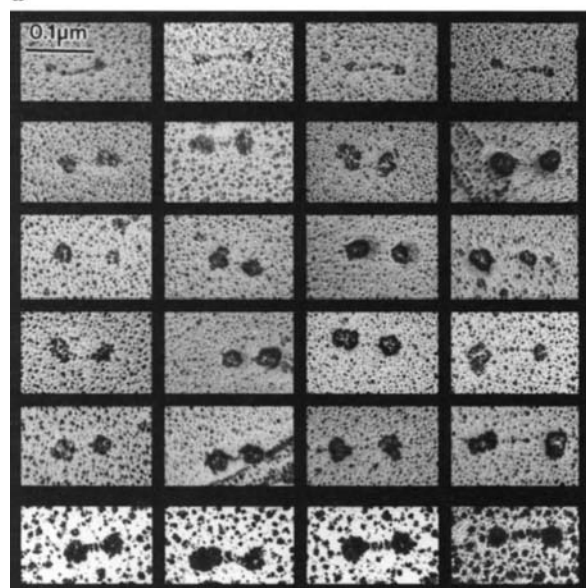
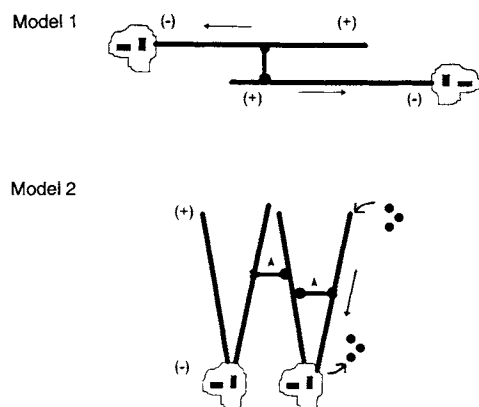


FIG. 3 KRP₁₃₀ molecules are bipolar structures capable of crosslinking microtubules. *a*, *Drosophila* KRP₁₃₀ is probably a homologue of *Xenopus* Eg5, a member of the BimC subfamily, consisting of 1,060 amino acid residues arranged in a tripartite motor–rod–tail manner, according to the map (top). The model (bottom) shows how four KRP₁₃₀ polypeptides could be arranged to produce bipolar homotetramers. Two parallel KRP₁₃₀ dimers assemble in an antiparallel fashion to form bipolar tetramers 96 nm in length with two 10-nm motor-domain 'heads' protruding on each end. The 'tails' of the KRP₁₃₀ subunits would be juxtaposed to the 'heads'; close packing of two heads and two tails would produce a structure visible in rotary-shadowed specimens as a single, globular domain approximately 20 nm × 20 nm on each end of the 60-nm rod. This is consistent with the dimensions of the rotary-shadowed KRP₁₃₀ molecules. *b*, Negatively stained images of complexes of KRP₁₃₀ with MTs. Sometimes the 'dumbbell-shaped' KRP₁₃₀ molecules bind by opposite ends to adjacent MTs (right), producing MT–MT crossbridges, whereas other molecules bind to MTs with one end only, producing projections from the MT wall, *c*, Individual fine, aperiodic crossbridges are observed in rotary-shadowed KRP₁₃₀–MT pellets (arrows). The reduced length of the crossbridges (53 nm) compared to the 'rods' of individual molecules (60 nm) may be due to occlusion of part of the 'rod' domain by MTs. Crossbridges were not observed in preparations of MTs alone (top row).

METHODS. Rotary shadowing was performed as described (Fig. 2 legend). The aperiodic crossbridges are difficult to shadow because they are obscured by MTs during platinum spraying. For negative staining, KRP₁₃₀–MT complexes were resuspended in buffer, loaded onto formvar-coated grids, then stained with 2% aqueous uranyl acetate for 3 min.

a**b**

them relative to one another during muscle contraction and cytokinesis²²⁻²⁴. Similarly, we propose that KRP₁₃₀ crosslinks adjacent MTs and slides them relative to one another. We propose that the bipolar KRP₁₃₀ tetramer binds to two antiparallel MTs emanating from opposite spindle poles and, by walking towards their plus ends, pushes the poles apart (Fig. 4b, model 1). This model is supported by the observation that several antibodies to members of the BimC subfamily react with KRP₁₃₀ and stain mitotic spindles, and by genetic studies which suggest that mem-

bers of the BimC subfamily participate in separation of the spindle poles^{3,6-11}. However, our data do not exclude the possibility that KRP₁₃₀ drives other forms of MT motility. For example, by crosslinking parallel MTs emanating from one or both poles it could drive MT flux and organize spindle poles⁵ or kinetochore fibres (Fig. 4b, model 2). It will now be interesting to test these models and determine whether other motors assemble into bipolar structures capable of driving MT–MT sliding in the mitotic spindle^{3-12,14,25,26} or interphase cytoplasm²⁷.

METHODS. Purified KRP₁₃₀ from the sucrose gradients was mixed with MTs and anti-Eg5 motor-domain antibody (Fig. 1) in amounts approximately equimolar to the amount of KRP₁₃₀. The MT–KRP₁₃₀–antibody complexes were pelleted through a glycerol cushion and processed for rotary shadowing as described (Fig. 2 legend).

Received 4 August; accepted 21 November 1995.

- Mitchison, T. J. *Phil. Trans. R. Soc. Lond.* **B336**, 99–106 (1992).
- Vernos, I. & Karsenti, E. *Trends Cell Biol.* **5**, 297–301 (1995).
- Enos, A. P. & Morris, N. R. *Cell* **60**, 1019–1027 (1990).
- LeGuellec, R., Paris, J., Couturier, A., Roghi, C. & Philippe, M. *Molec. cell. Biol.* **11**, 3395–3398 (1991).
- Sawin, K. E., LeGuellec, K., Philippe, M. & Mitchison, T. J. *Nature* **359**, 540–543 (1992).
- Roof, D. M., Meluh, P. B. & Rose, M. D. *J. Cell Biol.* **118**, 95–108 (1992).
- Hoyt, M. A., He, L., Loo, K. K. & Saunders, W. S. *J. Cell Biol.* **118**, 109–120 (1992).
- Hagan, I. & Yanagida, M. *Nature* **347**, 563–566 (1990).
- Hagan, I. & Yanagida, M. *Nature* **356**, 74–76 (1992).
- Sanders, W. S. & Hoyt, M. A. *Cell* **70**, 451–458 (1992).
- Heck, M. M. S. *et al. J. Cell Biol.* **123**, 665–679 (1993).
- Houliston, E., LeGuellec, R., Kress, M., Philippe, M. & LeGuellec, K. *Dev. Biol.* **164**, 147–159 (1994).
- Cole, D. G., Saxton, W. M., Sheehar, K. B. & Scholey, J. M. *J. biol. Chem.* **269**, 22913–22916 (1994).
- Barton, N. R., Pereira, A. J. & Goldstein, L. S. B. *Molec. Biol. Cell* **6**, 1563–1574 (1995).
- Huxley, H. E. *J. molec. Biol.* **7**, 281–308 (1963).

- Amos, L. A. *J. Cell Sci.* **87**, 105–111 (1987).
- Scholey, J. M., Heuser, J. A., Yang, J. T. & Goldstein, L. S. B. *Nature* **338**, 355–357 (1989).
- Hirokawa, N. *et al. Cell* **56**, 867–878 (1989).
- Tsuprun, V. L., Kashina, A. S. & Gelfand, V. I. *Dokl. Akad. Nauk SSSR* **320**, 1270–1272 (1991).
- Andrews, S. B., Gallant, P. E., Leapman, R. D., Schnapp, B. J. & Reese, T. S. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6503–6507 (1993).
- Sawin, K. E. & Mitchison, T. J. *Proc. natn. Acad. Sci. U.S.A.* **92**, 4289–4293 (1995).
- Huxley, A. F. & Nieredgerke, R. M. *Nature* **173**, 971–973 (1954).
- Huxley, H. E. & Henson, J. *Nature* **173**, 973–976 (1954).
- Yumura, S. & Fukui, Y. *Nature* **314**, 194–196 (1985).
- Hogan, C. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **90**, 6611–6615 (1993).
- Nislow, C., Lombillo, V. A., Kuriyama, R. & McIntosh, J. R. *Nature* **359**, 543–547 (1992).
- Koonce, M. P., Tong, J., Euteneuer, U. & Schliwa, M. *Nature* **328**, 737–739 (1987).
- Tyler, J. M. & Branton, D. *J. Ultrastruct. Res.* **71**, 95–102 (1980).

ACKNOWLEDGEMENTS. We thank R. S. Hawley and F. McNally for discussion, and R. Harris and K. Sheehan for technical assistance. This work was supported by grants from the March of Dimes Birth Defects Foundation (J.M.S.), the NIH (J.M.S. and W.M.S.), the Whitaker foundation (R.J.B.), and the American Cancer Society (J.M.S.). W.M.S. is an established investigator of the American Heart Association.

Post-transcriptional transactivation of human retroviral envelope glycoprotein expression by herpes simplex virus Us11 protein

Jean-Jacques Diaz, Madeleine Duc Dodon, Nathalie Schaerer-Uthurralt, Denis Simonin, Karine Kindbeiter, Louis Gazzolo & Jean-Jacques Madjar

Immuno-Virologie Moléculaire et Cellulaire, Université Claude Bernard Lyon-1/CNRS UMR30, Faculté de Médecine, rue Guillaume Paradin, 69372 Lyon CEDEX 08, France

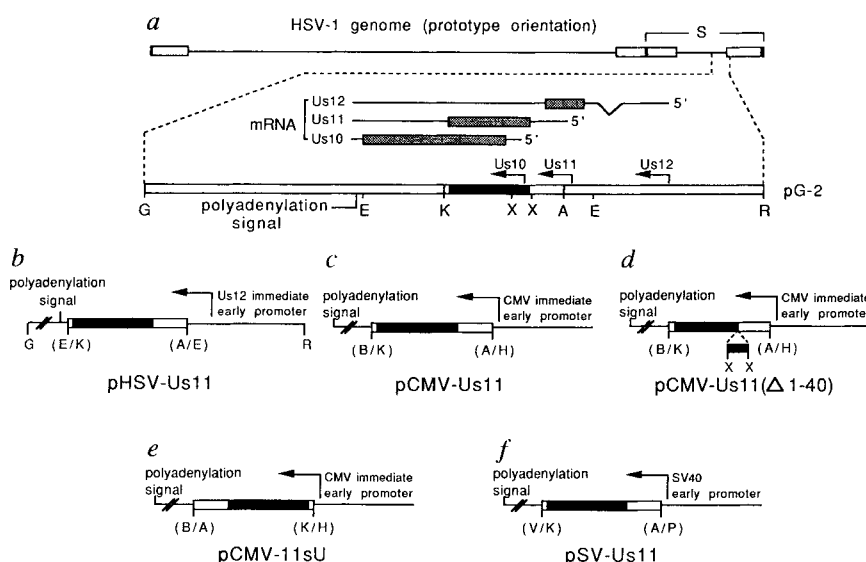
HERPES simplex virus type 1 (HSV-1) Us11 protein, a true late gene product packaged within the virion, is delivered into cells after infection, exhibits a nucleocytoplasmic localization at early times, and later accumulates in the nucleoli¹⁻⁵. This RNA-binding basic phosphoprotein, capable of oligomerization, is supposed to be involved in post-transcriptional regulation of gene expression after HSV-1 infection^{6,7}. Expression of human T-cell leukaemia/lymphoma virus type-I (HTLV-I) and of human immunodeficiency virus type 1 (HIV-1) is post-transcriptionally regulated by Rex and Rev, respectively⁸. These proteins are required for the cytoplasmic expression of unspliced *gag-pol* and singly spliced *env* transcripts^{9,10}. Here we show that HSV-1 Us11 protein is able to bind Rex- and Rev-responsive elements and to transactivate envelope retroviral glycoprotein expression.

To determine whether Us11 protein could substitute for HTLV-I Rex protein, HeLa cells were cotransfected with three different vectors; the first expressing HTLV-I envelope (*env*) glycoprotein through a non-spliceable primary transcript; the second expressing Tax transcriptional transactivator (*tax*); and

the third expressing either Rex or Us11 protein (Fig. 1 and legend to Fig. 2). Controls included HeLa cells cotransfected with either the *env* and *tax* vectors only, or with the same two vectors together with non-functional Us11-derived constructs (pG-2 and pCMV-11sU depicted in Fig. 1). After transfection, expression of functional envelope glycoproteins was assayed by scoring multinucleated cells containing at least five nuclei. As expected, expression of *rex* together with *env* and *tax* led to the formation of syncytia, whereas in control cultures no such multinucleated cells were observed; these multinucleated cells were also obtained under expression of Us11 gene (Fig. 2a) and in a dose-dependent manner (data not shown). However, the number of syncytia was about 25–50-fold lower under expression of Us11 than under that of *rex* for the same amount of transfected plasmid. Immunofluorescence analyses showed Us11 protein concentrated in the nucleoli of each multinucleated cell, but also present in their cytoplasm (Fig. 2b), whereas Tax was nuclear and excluded from nucleoli (Fig. 2c). Exposure of cells to a neutralizing monoclonal antibody directed against the envelope glycoprotein, resulted in a 75% decrease in the number of syncytia (not shown, see legend to Fig. 2a), indicating that formation of multinucleated cells was due to the expression of HTLV-I envelope glycoprotein. In addition, synthesis and cleavage of HTLV-I envelope glycoprotein precursor induced by Us11 expression was confirmed by western blot using an antiserum directed against the transmembrane gp21 (Fig. 2d). Deletion of the first 40 amino acids of Us11 protein abolished the Us11-induced syncytia formation (pCMV-Us11 (Δ 1–40) in Fig. 2a) even though the truncated form of Us11 protein could be detected by western blot and by immunofluorescence concentrated in the nucleoli (not shown).

To verify whether Us11 protein could also function as HIV-1 Rev protein, HeLa cells were cotransfected with three different vectors; the first expressing HIV-1 envelope (*env*) glycoprotein through a non-spliceable primary transcript; the second expressing Tat transcriptional transactivator (*tat*); and the third expressing either Rev or Rex or Us11 protein. Controls included HeLa cells cotransfected with either the *env* and *tat* vectors only, or with the same two vectors together with non-functional Us11-derived construct (pCMV-11sU depicted in Fig. 1). Seventeen hours

FIG. 1 Expression vectors. **a**, Construction of pG-2. The *Bgl*II–*Eco*RI fragment of the short (S) region of HSV-1 was subcloned into pBluescript (Stratagene)⁷. Bent arrows, transcription start sites for Us10, Us11 and Us12 primary transcripts. The position of the common polyadenylation signal is specified. The 3' coterminal family of mRNA expressed from this region is shown above the corresponding coding sequences¹⁸. Hatched boxes, regions translated in the different mRNA. In this pG-2 construct, Us11 coding sequence is under the control of its own true late promoter. Under these conditions, Us11 protein is not expressed after transfection of pG-2, as verified by western blot analyses (not shown). **b**, Construction of pHSV-Us11. The pG-2 *Ksp*I–*Apa*LI fragment containing the Us11 coding sequence (black box) was subcloned in pG-2 between the two *Eco*NI restriction sites. The Us11 sequence is under the control of the Us12 immediate-early promoter. **c**, Construction of pCMV-Us11. The same pG-2 *Ksp*I–*Apa*LI fragment was subcloned between the *Bam*HI–*Hind*III restriction sites of pBC12/CMV¹⁹, with the Us11 sequence under the control of the CMV immediate-early promoter. The polyadenylation signal was brought by pBC12/CMV. **d**, Construction of pCMV-Us11 (Δ 1–40). The *Xho*I–*Xho*I fragment was removed from pCMV-Us11. The coding sequence of Us11 protein contains three in-frame AUG codons¹⁸, the first two being located between these two *Xho*I restriction sites. pCMV-Us11 (Δ 1–40) gives rise to an Us11 protein without its first 40 amino acids. **e**, Construction of pCMV-11sU. Constructed as pCMV-Us11 except that the *Ksp*I–*Apa*LI was inserted in the reversed orientation in pBC12/CMV. **f**,



Construction of pSV-Us11. The *Ksp*I–*Apa*LI fragment from pG-2 was inserted between the *Xba*I–*Pst*I restriction sites of pSVK3 (Pharmacia). The Us11 coding sequence was under the control of the SV40 early promoter and the polyadenylation signal, brought by pSVK3, was from the SV40 early gene. Endonuclease restriction sites: G, *Bgl*II; E, *Eco*NI; K, *Ksp*I; A, *Apa*LI; R, *Eco*RI; B, *Bam*HI; H, *Hind*III; X, *Xho*I; V, *Xba*I; P, *Pst*I. The endonuclease restriction sites that were destroyed during the cloning are in parentheses.

later, these transfected cells were cocultivated with CD4-positive HeLa cells expressing the β -galactosidase gene under the control of the HIV-1 long terminal repeat (LTR). Expression of functional envelope glycoproteins was assayed by estimating the number of β -galactosidase-positive cells obtained via *rev*, *rex* or Us11 gene expression. Under these experimental conditions (Fig. 3a) Rex as well as Us11 protein induced formation of multinucleated cells expressing β -galactosidase, to a similar extent, albeit about five to six times less efficiently than Rev. These results indicate that Rex and Us11 protein exhibit an equal efficiency in transactivating expression of a HIV-1 non-spliceable envelope messenger RNA. Taken together these results demonstrate that HSV-1 Us11 protein can function as both Rex and Rev in transactivating expression of non-spliceable envelope mRNA, in transient expression assays.

The ability of Us11 to induce, like Rex and Rev^{11,12}, an accumulation of unspliced envelope mRNA in the cytoplasm was then investigated by using vectors giving rise to spliceable primary transcripts. No cytoplasmic expression of HIV-1 unspliced mRNA¹² could be obtained under coexpression of Us11 protein, nor was it possible to rescue HIV-1 expression from an infectious molecular clone deficient in the *rev* gene (not shown). Conversely, coexpression of Us11 protein in cultures transfected with pgTAX-LTR¹³ resulted in the cytoplasmic appearance of an additional mRNA corresponding to the unspliced HTLV-I *env* transcript (Fig. 3b, compare lane 1 with lanes 2 and 3) and in syncytia formation in a dose-dependent

manner (Fig. 3c). In these transfected cells, the unspliced transcripts of β -actin was detected exclusively in the nucleus and never in the cytoplasm (Fig. 3d). This indicates that the presence of the unspliced *env* mRNA in the cytoplasm, found under expression of Us11, is not due to a non-specific RNA efflux from the nucleus or to a non-specific inhibition of cellular splicing. These results suggest that Us11 protein may transactivate envelope glycoprotein expression through a mechanism similar to that of Rex and Rev, by interacting with the Rex- and Rev-responsive elements (XRE and RRE). Because the binding of Us11 protein to RNA is specific to sequence and secondary structure for at least two HSV-1 mRNAs^{6,14}, we investigated the ability of Us11 protein to interact specifically with the XRE and also with the RRE. This has been done by gel-retardation assays with a GST-Us11 fusion protein and different ³²P-labelled RNA target sequences. A part of the HSV-1 UL34 mRNA, encompassing the binding site of Us11 protein⁶ was used as a positive control (Fig. 4a). As expected^{15,16}, GST-Rev and GST-Rex proteins bind their respective responsive elements (Fig. 4b,c). GST-Rex is also able to bind the RRE¹⁷ (Fig. 4b), whereas GST-Rev does not bind the XRE¹⁷ (Fig. 4c). Conversely, GST-Us11 protein binds the RRE (Fig. 4b) and even better the XRE (Fig. 4c) giving rise to at least two distinct RNA-protein complexes with retarded mobility. These RNA-protein complexes were observed in the presence of up to 10 μ g yeast transfer RNA, whereas they were inhibited by addition of specific competitor unlabelled XRE or RRE transcripts (not shown). In the same conditions GST-Us11

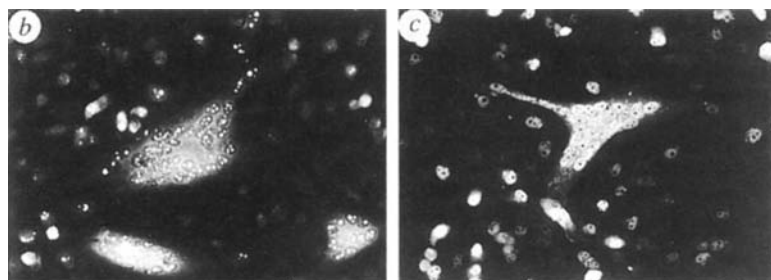
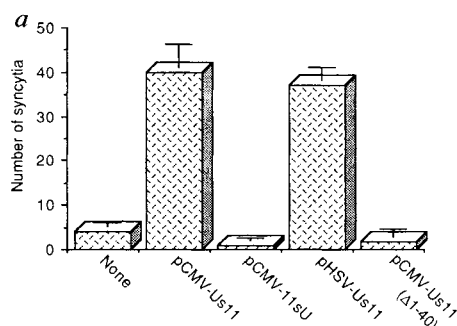
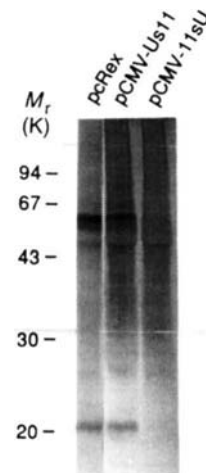


FIG. 2 Expression of functional HTLV-I envelope glycoprotein. *a*, Syncytia formation under coexpression of Tax-directed HTLV-*env* and Us11 protein. In every case, 18×10^4 HeLa cells, in 35-mm diameter Petri dishes, were cotransfected with 1 μ g pHTLV-*env* and 1 μ g pTAX²⁰ expression vectors only (none) or with these two vectors together with 1 μ g pCMV-Us11 or pHSV-Us11 or pCMV-Us11 (Δ 1-40), using the calcium phosphate method²¹. The pHTLV-*env* vector was constructed by replacing the *Hind*III—*Stu*I sequence of pHTLV-tat1 vector²² between the two HTLV-I long terminal repeats, with the *Hind*III—*Pst*I fragment containing the HTLV-I envelope coding sequence. The *tax* and *rex* coding sequences were placed under the control of the CMV promoter in pTAX²⁰ and in pCRex²⁰ vectors. In control experiments, a non-functional Us11-derived construct, pCMV-11sU, was cotransfected with the *env* and *tax* expression vectors. The number of syncytia containing at least 5 nuclei was scored under an inverted microscope at low magnification ($\times 20$) after May-Grünwald-Giemsa staining. Data are shown by the mean $\pm$ s.e. from three determinations. These cotransfection experiments were done with an identical amount of total DNA (4 μ g), by supplementation of the different vectors with pUC18 DNA. Inhibition of syncytia formation was undertaken by addition to the culture medium of a monoclonal antibody (LAT27) directed against an epitope of the external part of HTLV-I envelope glycoprotein. This monoclonal antibody, able to neutralize the HTLV-I-mediated transformation of normal peripheral blood T lymphocytes²³, decreased both the Us11 and the Rex-induced syncytia formation. No inhibition was observed in cultures treated with another monoclonal antibody (LAT12 devoid of neutralizing activity) directed against HTLV-I envelope glycoprotein. *b*, *c*, Expression of Us11 (*b*) and of Tax (*c*) proteins by immunofluorescence analyses. Forty-eight hours after transfection, cells were fixed with paraformaldehyde, permeabilized with Triton X-100 and localization of Us11 and Tax proteins was assayed with anti-Us11 (ref. 7) or anti-Tax (provided by B. R. Cullen)

rabbit antibodies and fluorescein isothiocyanate-labelled second antibody. *d*, Western blot analysis of envelope glycoproteins. HeLa cells (4.5×10^6) were cotransfected with the Tax-directed HTLV-*env* expression vector together with either pCRex or pCMV-Us11 or pCMV-11sU. Forty hours after transfection, cells were washed and lysed in 0.02 M Tris-HCl pH 8.0, 0.12 M NaCl containing 0.2 mM phenylmethylsulphonylfluoride, aprotinin (5 μ g ml⁻¹), 0.2 mM EGTA, 0.2 mM NaF, 0.2% sodium deoxycholate and 0.5% Nonidet P-40. Lysates were cleared for 1 h at 80,000g, then incubated for 4 h with lentil-lectin-Sepharose (Pharmacia) at 4 °C. The lectin-bound proteins were eluted with 0.2 M α -methyl-mannoside, then immunoprecipitated using a serum from a patient with HTLV-I-associated myelopathy. The immunoprecipitates were then isolated on Pansorbin (Calbiochem), separated by electrophoresis on a 10% polyacrylamide gel in the presence of SDS and blotted onto a nitrocellulose membrane. Envelope glycoproteins (gp63 and gp21) were revealed using the SP-11 rabbit anti-serum²⁴.



protein was unable to interact with its own mRNA as well as with a non-specific RNA transcript (Fig. 4d). To verify whether transactivation of HTLV-I envelope expression was the consequence of this specific interaction of Us11 protein with the XRE, the sequence coding for the XRE was removed from pHTLV-*env* expression vector (depicted in Fig. 2a). No expression of envelope glycoprotein could be detected under coexpression of either Rex or Us11 proteins (Fig. 4e).

This report provides, to our knowledge, the first evidence that the product of an heterologous virus, Us11 protein of HSV-1, can intervene post-transcriptionally in the life cycle of human retro-

viruses. This was clearly demonstrated for both retroviruses when unsplicable primary transcripts were expressed, and only for HTLV-I when *env* mRNA were derived from spliceable primary transcripts. Because Us11 protein is able to interact directly with the XRE and with the RRE, it is very likely that this transactivation is mediated by a mechanism similar to that shared by Rex and Rev. This is further supported by the lack of such a transactivation when the XRE is removed from the HTLV-I envelope mRNA. Furthermore, these data raise the possibility of *in vivo* interactions between herpes virus and human retroviruses, possibly affecting expression of each virus at the post-transcriptional level. □

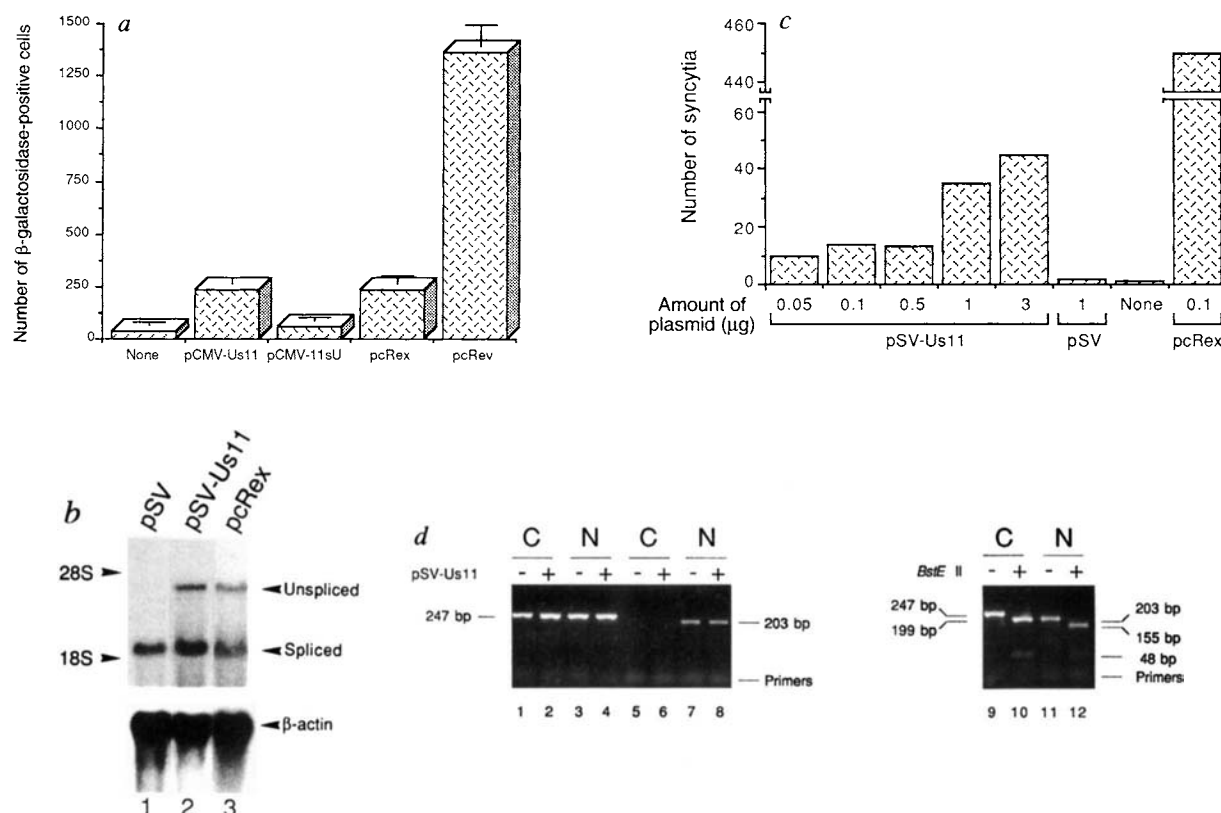
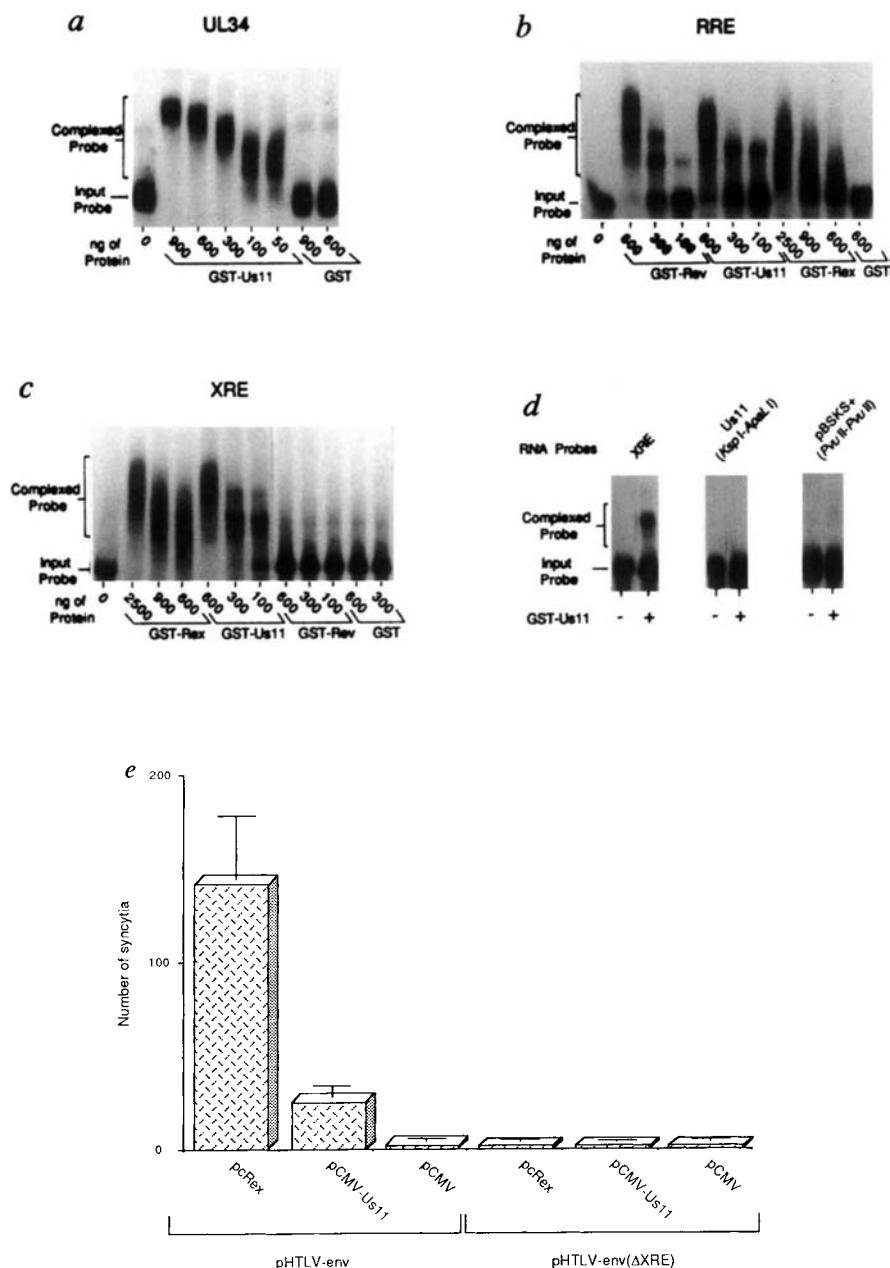


FIG. 3 Syncytia formation under coexpression of Tat-directed HIV-*env* and either Us11, or Rex, or Rev, as well as Us11-induced HTLV-I envelope glycoprotein expression through a non spliced mRNA. **a**, Syncytia formation was scored after cocultivation of H12 indicator cells²⁵ (HeLa CD4-positive cells carrying the bacterial LacZ gene under control of HIV-1 LTR) and HeLa cells transfected with pLH₃²⁶ (expressing HIV-1 envelope glycoprotein under control of HIV-2 LTR) and pcTAT¹², together with 1 μ g of either pCMV-Us11 or pcRex or pcRev. HeLa cells were cotransfected²¹ with 1 μ g pLH₃ and 1 μ g pcTAT expression vectors as described for HTLV-I in the legend to Fig. 2. After 17 h, cells were collected, counted then cocultivated with H12 indicator cells. After 24 h, cells were fixed and assayed for β -galactosidase activity. All blue foci were scored at low magnification ($\times 20$). Means $\pm$ s.e. for three experiments. **b**, Northern blot analysis of cytoplasmic HTLV-I envelope message. HeLa cells (7.5×10^5) were cotransfected with pgTAX-LTR with either pSV (pSVK3), or pSV-Us11 (Fig. 1) or pcRex. Total cytoplasmic RNA was purified by the SDS-proteinase K method. For electrophoresis, each lane contained 8 μ g total cytoplasmic RNA. Blotting and hybridization used standard protocols and *env* mRNA was detected using the DNA fragment coding for the XRE as a probe, after ³²P-labelling by random priming. The level of cellular mRNA expression was monitored by visualizing the β -actin mRNA revealed by hybridization with a ³²P-labelled complete mouse β -actin coding sequence²⁷. **c**, Syncytia formation under HTLV-I envelope glycoprotein expression. HeLa cells (10^5) in 35-mm diameter Petri dishes were cotransfected²¹ with 50 ng pgTAX-LTR¹³ vector and either pSV (pSVK3), or increasing amounts of pSV-Us11, or pcRex as positive control. Numbering of syncytia was as described for Fig. 2a but at a

higher magnification ($\times 100$). The Rex-responsive expression vector pgTAX-LTR contains the two coding exons for Tax and Rex, the complete *env* gene and the entire 3'LTR, inserted into the pBC12/CMV¹⁹ vector. A mutation introduced at the translational initiation codon of Rex prevents its synthesis from pgTAX-LTR. **d**, Subcellular distribution of spliced and unspliced β -actin transcripts under coexpression of HTLV-I envelope glycoprotein and Us11 protein. The subcellular distribution of β -actin transcripts was determined by RT-PCR²⁸. Total nuclear RNA purified from pelleted nuclei²⁹ and total cytoplasmic RNA were treated with 10 U DNase I, RNase free (Boehringer) for 1 h at 20 °C. Total nuclear (N) or total cytoplasmic (C) RNA (5 μ g) was reverse transcribed at 37 °C for 1 h with 200 U SuperScript RNase H⁻ (Life Technologies) and 1 μ g oligo dT₁₂₋₁₈ (Sigma). For amplification of cDNA of spliced actin mRNA, primers ACTINS and ACTINA²⁸ were used with 1/10th of the cDNA products (**d**, lanes 1–4 and 9). For amplification of cDNA of unspliced actin primary transcript, primers ACTINS-1 and ACTINA²⁸ were used with 1/10th of the cDNA products (**d**, lanes 5–8 and 11). Of the PCR products, 1/5th was submitted to electrophoresis in a non-denaturing gel, made of 0.5% agarose and 1.5% Infinity Agarose Enhancer (Oncor) containing 0.5 μ g ml⁻¹ ethidium bromide. PCR amplifications with primers ACTINS and ACTINA and with primers ACTINS-1 and ACTINA give rise to DNA fragments of 247 bp and of 203 bp, respectively. No amplification was obtained with RNA not previously reverse-transcribed, as control (not shown). Digestion with BstEII of the 247 bp DNA fragment resulted in the appearance of two DNA fragments of 199 bp and 48 bp (**d**, lane 10). Digestion of the 203 bp DNA fragment resulted in the appearance of two DNA fragments of 155 bp and 48 bp (**d**, lane 12), as expected³⁰.

FIG. 4 Binding properties of GST-Usl1 protein to the XRE and RRE, and absence of syncytia formation after removal of the XRE. **a**, Binding properties of GST-Usl1 protein to UL34 RNA by gel-retardation assay. UL34 cDNA encoding the binding site of Usl1 protein⁶ was cloned into plasmid vector pBluescript (Stratagene) using the *Sma*I and *Eco*RI restriction sites of the vector, and the *Eco*NI and *Eco*RI restriction sites of UL34 cDNA. UL34 RNA was synthesized *in vitro* using T3 RNA polymerase (Promega) according to the supplier's protocol, except that the UTP concentration was decreased to 100 μ M and that [α -³²P]UTP (Amersham) was included at 800 Ci mmol⁻¹. Reactions were terminated by DNase digestion followed by extractions with phenol-chloroform. RNA was recovered after ethanol precipitation. The entire Usl1 coding sequence was amplified by PCR and cloned in frame with the coding sequence of glutathione S-transferase in pGEX-2T (Pharmacia). The resulting plasmid, pGEX-Usl1, was maintained in *E. coli* strain DH5 α (Life Technologies) and the recombinant protein expressed by induction of log-phase cells with 0.1 mM isopropyl- β -D thiogalactoside. Cells were lysed by sonication in PBS contained 1% Triton X-100. GST-Usl1 was purified from the lysate by passage over a glutathione Sepharose 4B (Pharmacia) affinity column, followed by elution with 10 mM reduced glutathione (Sigma). After dialysis and concentration (15 mM Tris-HCl pH 6.8) the amount of protein was determined by the method of Bradford. Purity and integrity of the recombinant protein were assessed by SDS-PAGE followed by Coomassie blue staining. A single band was obtained for GST-Usl1. Gel retardation assay: samples containing GST-Usl1 or GST at the given concentrations and UL34 RNA (previously denatured at 80 °C for 5 min and renatured at 37 °C for 20 min) were incubated 10 min at 4 °C, then submitted to electrophoresis at 4 °C on a standard TBE 4% polyacrylamide gel containing 4% glycerol for 450 Vxh. All samples (20 μ l) contained 500 ng yeast tRNA (Boehringer), and 2×10^5 c.p.m. (~ 0.8 ng) of the ³²P-internally labelled UL34 RNA (480 bases). Gels were dried and submitted to autoradiography for 45 min. **b**, Comparison of the RRE-binding properties of GST-Rex, GST-Usl1 and GST-Rex proteins. The cDNA encoding the RRE was amplified by PCR from pLH₃ and cloned into plasmid vector pBluescript using the *Sma*I restriction site of the vector. RNA containing the RRE was prepared as described above for UL34 RNA except that T7 instead of T3 RNA polymerase was used. The entire Rex and Rev coding sequences were amplified by PCR from pcRex and pcRev, respectively, and cloned in frame with the coding sequence of GST as above for GST-Usl1. The resulting plasmids, pGEX-Rex and pGEX-Rev, were maintained in *E. coli* strain DH5 α and the recombinant proteins were purified as above for GST-Usl1. After SDS-PAGE and western blot, several bands were obtained from the GST-Rex preparation, the full-length fusion protein representing less than 10%, whereas only a single band was obtained for GST-Rev. Gel retardation assays were performed with samples containing either GST-Rev, or GST-Usl1, or GST-Rex, or GST proteins at the given concentrations and the RRE (denatured and renatured as above). All samples contained 500 ng yeast tRNA, and 2×10^5 c.p.m. (about 1 ng) of the ³²P-internally labelled RRE (234 bases). **c**, Comparison of the XRE-binding properties of GST-Rex, GST-Usl1 and GST-Rex proteins. The cDNA encoding the XRE was amplified by PCR from pgTAX-LTR and cloned into plasmid vector pBluescript as above for the cDNA encoding the RRE. RNA containing the XRE was prepared as described above for UL34 RNA. Gel retardation assays used samples containing either GST-Rex, or GST-Usl1, or GST-Rex, or GST proteins at the given concentrations and the XRE (denatured and renatured as above).



All samples contained 500 ng yeast tRNA, and 2×10^5 c.p.m. (~ 1 ng) of the ³²P-internally labelled XRE (285 bases). **d**, Binding of GST-Usl1 protein to the XRE, Usl1 RNA and a nonspecific transcript. Gel retardation assays used samples containing 100 ng GST-Usl1 and either the XRE or a part of Usl1 mRNA or an RNA transcript obtained from pBluescript (pBSKS+). A part of Usl1 mRNA was synthesized from the *Ksp*I-*Apa*LI (Fig. 1a) fragment of Usl1 cDNA. The cDNA encoding the nonspecific RNA corresponded to the *Pvu*II-*Pvu*II fragment of the pBluescript vector. The XRE and nonspecific RNA were prepared as described above for UL34 RNA. All samples contained 5 μ g yeast tRNA, and 2×10^5 c.p.m. of the ³²P-internally labelled RNA. **e**, HeLa cells (15×10^4) in 35-mm diameter Petri dishes were cotransfected²¹ with either 1 μ g pHTLV-env or 1 μ g pHTLV-env(Δ XRE), together with 1 μ g pcTAX and either 1 μ g pcMV-Usl1 or 0.5 μ g pcRex. The total amount of DNA was adjusted to 3 μ g by supplementation with pcMV. The pHTLV-env(Δ XRE) vector was constructed by removing from pHTLV-env the *Nde*I-*Bgl*II DNA fragment (~ 600 bp) containing the sequence coding for the XRE. Expression of the modified env mRNA was verified by RNase protection assay (not shown). pcMV is pBC12/CMV¹⁹. Numbering of syncytia was as described for Fig. 2a but at a higher magnification ($\times 200$). Mean $\pm$ s.e. from four determinations.

Received 18 October; accepted 6 November 1995.

1. Massé, T., Garcin, D., Jacquemont, B. & Madjar, J.-J. *Molec. gen. Genet.* **220**, 377–388 (1990).
2. Roller, R. J. & Roizman, B. *J. Virol.* **66**, 3624–3632 (1992).
3. MacLean, C. A., Rixon, F. J. & Marsden, H. S. *J. gen. Virol.* **68**, 1921–1937 (1987).
4. Puvion-Dutilleul, F. *Eur. J. Cell Biol.* **43**, 487–498 (1987).
5. Johnson, P. A., MacLean, C., Marsden, H. S., Dalziel, R. G. & Everett, R. D. *J. gen. Virol.* **67**, 871–883 (1986).
6. Roller, R. J. & Roizman, B. *J. Virol.* **65**, 5873–5879 (1991).
7. Diaz, J.-J. *et al. J. gen. Virol.* **74**, 397–406 (1993).
8. Cullen, B. R. *Microbiol. Rev.* **56**, 375–394 (1992).
9. Malim, M. H., Hauber, J., Le, S.-Y., Maize, J. V. & Cullen, B. R. *Nature* **338**, 254–257 (1989).
10. Ballaun, C. *et al. J. Virol.* **65**, 4408–4413 (1991).
11. Inoue, J.-I., Itoh, M., Akizawa, T., Toyoshima, H. & Yoshida, M., *Oncogene* **6**, 1753–1757 (1991).
12. Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. *Nature* **335**, 181–183 (1988).
13. Rimsky, L., Duc Dodon, M., Dixon, E. P. & Greene, W. C. *Nature* **341**, 453–456 (1989).
14. Roller, R. J. & Roizman, B. *J. Virol.* **64**, 3463–3470 (1990).
15. Malim, M. H. & Cullen, B. R. *Cell* **65**, 241–248 (1991).
16. Bogerd, H. P., Tiley, L. & Cullen, B. R. *J. Virol.* **66**, 7572–7575 (1992).
17. Bogerd, H. P., Huckaby, G. L., Ahmed, Y. F., Hanly, S. M. & Greene, W. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5704–5708 (1991).

18. Rixon, F. J. & McGeoch, D. J. *Nucleic Acids Res.* **12**, 2473–2487 (1984).
19. Cullen, B. R. *Cell* **46**, 973–982 (1986).
20. Rimsky, L. *et al. Nature* **335**, 738–740 (1988).
21. Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
22. Nerenberg, M., Hinrichs, S. H., Reynolds, R. K., Khoury, G. & Jay, G. *Science* **237**, 1324–1329 (1987).
23. Tanaka, Y., Zeng, L., Shitaki, H., Shida, H. & Tozawa, H. *J. Immun.* **147**, 354–360 (1991).
24. Palker, T. J. *et al. J. Virol.* **66**, 5879–5889 (1992).
25. Chameau, P., Alizon, M. & Clavel, F. *J. Virol.* **66**, 2814–2820 (1992).
26. Emerman, M., Vazeux, R. & Peden, K. *Cell* **57**, 1155–1165 (1989).
27. Hanahan, D. J. *molec. Biol.* **166**, 557–580 (1983).
28. Tan, W. & Schwartz, S. J. *Virol.* **69**, 2932–2945 (1995).
29. Stallcup, M. R. & Washington, L. D. *J. biol. Chem.* **258**, 2802–2807 (1993).
30. Nakajima-Iijima, S., Hamada, H., Reddy, P. & Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6133–6137 (1985).

ACKNOWLEDGEMENTS. J.-J.D. and M.D.D. contributed equally to this work. We are grateful to M. Alizon, B. R. Cullen, M. Emerman, W. C. Greene and G. Jay who kindly provided different cells and constructs and to Y. Tanaka for his generous gift of LAT27 antibody. We are grateful to R. Grantham and B. B. Rudkin for critical reading of the manuscript. This work was supported by grants to L.G. and to J.-J.M. from Agence Nationale de Recherches sur le SIDA (A.N.R.S.). A fellowship to D.S. from ANRS is gratefully acknowledged.

Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2

Michael J. Eck^{*†}, Scott Pluskey^{†‡}, Thomas Trüb^{†‡}, Stephen C. Harrison^{*§} & Steven E. Shoelson^{†‡}

§ Howard Hughes Medical Institute, * Laboratory of Molecular Medicine, Children's Hospital, ‡ Research Division, Joslin Diabetes Center, and † Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA

THE domain organization of many signalling proteins facilitates a segregation of binding, catalytic and regulatory functions^{1,2}. The mammalian SH2 domain protein tyrosine phosphatases (PTPs) contain tandem SH2 domains and a single carboxy-terminal catalytic domain³. SH-PTP1 (PTP1C, HCP) and SH-PTP2 (Syp, PTP2C, PTP1D) function downstream from tyrosine kinase-linked insulin, growth factor, cytokine and antigen receptors^{4–12}. As well as directing subcellular localization by binding to receptors and their substrates, the two SH2 domains of these PTPs function together to regulate catalysis^{7,13,14}. Here we report the structure of the tandem SH2 domains of SH-PTP2 in complex with monophosphopeptides. A fixed relative orientation of the two domains, stabilized by a disulphide bond and a small hydrophobic patch within the interface, separates the peptide binding sites by ~40 Å. The defined orientation of the SH2 domains in the structure, and data showing that peptide orientation and spacing between binding sites is critical for enzymatic activation, suggest that spatial constraints are important in this multidomain protein–protein interaction.

The structure of the tandem Src homology (SH2) domain portion of human SH-PTP2 (residues 1–225), complexed with two phosphotyrosyl peptides each with 11 residues, that correspond to the Tyr 1009 binding site of the platelet-derived growth factor (PDGF) receptor, has been refined to a crystallographic *R*-value of 17.2 at 2.1-Å resolution. The electron density at the interface between SH2 domains is shown in Fig. 1; details of the structure determination are described in the legend. Density for the 11 residue 1009 phosphotyrosyl peptides is clear in the binding pockets of both domains.

Both the amino- and carboxy-SH2 domains of SH-PTP2 adopt the conserved SH2 fold (Fig. 2). In the orientation shown, both phosphopeptide binding sites lie on the surface facing the viewer,

but they are widely separated and run almost antiparallel to one another (Fig. 2*b,c*). The N and C termini of the protein and the interdomain linker are on the opposite face of the molecule. The short polypeptide segment linking the two domains is well ordered and contributes to the interface between the domains. Interactions between domains involve primarily the A helix of the N-SH2 domain, the interdomain linker, and the βD' strand and D'E turn of the C-SH2 domain (Figs 1 and 2*b*). A disulphide bond links Cys 104, near the C terminus of the N-SH2 domain, with Cys 174 in the D' strand of the C-SH2 domain. The disulphide bond lies within the domain interface at one edge of a small hydrophobic core, which is continuous with the hydrophobic cores of both SH2 domains. In addition to Cys 104 and Cys 174, the interdomain hydrophobic core includes residues Leu 19, Leu 102, Leu 177 and the aliphatic portion of the side chain of Arg 23. A salt bridge between the side chains of Arg 5 and Asp 106 also links the two domains. Nearer the 'front' of the protein, as viewed in Fig. 2*b*, the domains are closely apposed but not in direct contact, creating a narrow crevice filled with ordered water molecules.

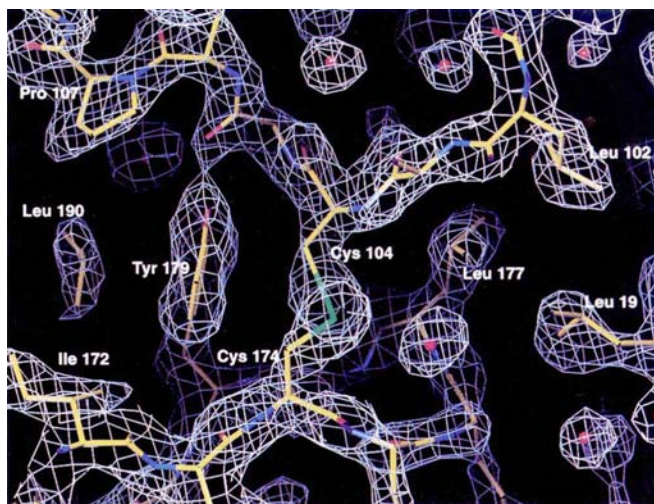
Disulphide bonds are uncommon in cytoplasmic proteins. The one in the structure presented here is largely buried within the hydrophobic core of the interdomain region, and is resistant to reduction (the crystals were grown in 10 mM DTT). The juxtaposition of the cysteines, which results from the structure surrounding them (Fig. 1), suggests that they would be prone to oxidation even in a cellular environment, but we cannot be certain that these cysteines are oxidized *in vivo*.

The interactions of the N-terminal SH-PTP2 domain with a series of phosphopeptides, including 1009 of the PDGF receptor, have been described in detail previously¹⁵. The structures of the N-SH2 domain within this tandem construct and the isolated N-SH2 domain are essentially interchangeable, as are their modes of 1009 peptide recognition. The N-SH2 and C-SH2 domains and their bound phosphopeptides are overlaid in Fig. 3. The two domains are 50% identical in primary sequence, and their structures superimpose well in the region of the A helix and central sheet. The comparison shows a significant difference between the relative positions of the B helix and D', E and F strands of the two domains. Nevertheless, the conformation of the 1009 peptide in complex with the two domains is nearly identical. Surprisingly, the greatest divergence in the structure of the bound phosphopeptides is observed for phosphotyrosine itself (Fig. 3).

The fixed orientation of the SH2 domain binding sites relative to one another has implications for phosphoprotein recognition. SH-PTP2 interacts in cells with IRS-1 (ref. 4) and the PDGF receptor^{5–8,16} following stimulation with insulin or PDGF, respectively. When phosphorylated, tyrosines 1172 and 1222 of IRS-1 associate selectively *in vitro* with the N- and C-SH2 domains, respectively^{13,14}. To probe orientational relationships between

FIG. 1 View of the refined atomic model and 2.1-Å resolution electron-density map in the interdomain region of the tandem domain. The $2F_o - F_c$ map is contoured at 1.3σ . The interdomain linker extends across the top of the figure (Leu 102 to Pro 107). Cys 104 in the linker forms a disulphide bond with Cys 174 in the D'E turn of the C-SH2 domain (Ile 172 to Tyr 179). Residues from the linker (Leu 102 and Cys 104) and both the N-SH2 (Leu 19) and C-SH2 domains (Cys 174 and Leu 177) create a hydrophobic region between the domains that is continuous with the hydrophobic cores of the domains themselves. Red spheres represent ordered solvent molecules. Several interactions stabilize the conformation of the interdomain linker in addition to the disulphide bond. Residues Leu 102 and Asn 103 immediately preceding Cys 104 are in essentially the same positions in the previously determined structure of the isolated N-SH2 domain¹⁵. The carbonyl of Ala 105 is coordinated by the Tyr 179 hydroxyl group, and Pro 107 and the methyl group of Thr 108 (not shown) pack into a hydrophobic surface created by Trp 112, Phe 113, Leu 190, Phe 147 and Tyr 179. His 8 (not shown) can form hydrogen bonds with the carbonyl of Leu 102 and the side chain of Asn 103.

METHODS. A fragment of the human SH-PTP2 cDNA was subcloned into the NcoI site of a pET-14b vector (Novagen) using the polymerase chain reaction (PCR). *Escherichia coli* strain BL21(DE3) was transformed with the expression vector encoding residues 1–225 of human SH-PTP2. Following isopropyl thio β -D-galactoside (IPTG) induction of protein expression and cell collection and lysis, the protein was isolated and purified by affinity chromatography using an immobilized phosphopeptide column prepared by linking the sequence surrounding Tyr 1222 of IRS-1 (LSTpYASINFQK) to Affi-Gel 10 (Bio-Rad). Bound protein was eluted from the washed beads with 25 mM HEPES, 6.0 M urea, 10 mM DTT, pH 7.4, and dialysed at 4 °C against 20 mM Tris-HCl, 100 mM NaCl, 10 mM DTT, pH 8.0, for > 16 h. The recovered protein was further purified by gel-filtration fast protein liquid chromatography (FPLC) using a Superdex-75 column equilibrated in dialysis buffer, concentrated in a CentrPrep device to 22 mg ml⁻¹, and combined with a fourfold molar excess of PDGF receptor phosphopeptide 1009, SVLPYTA VQNP²². Crystals were grown using the hanging drop vapour diffusion method by combining 2 μ l of protein + peptide solution with 2 μ l of a well solution containing 20–30% PEG 4,000, 200 mM sodium or ammonium acetate, 100 mM Tris, and 10 mM DTT, pH 8.5. The monoclinic crystals (space group $P2_1$, $a = 36.16$, $b = 48.48$, $c = 63.87$, $\alpha = \gamma = 90^\circ$, $\beta = 104.97^\circ$) grew as clusters of large, but very thin, plates over several days at 22 °C. Crystals were transferred to a synthetic mother liquor containing 10% glycerol as a cryoprotectant and flash-cooled in a boiling nitrogen stream at -160°C for data collection. Diffraction data from a single crystal were recorded with a 30-cm MarResearch image plate scanner using an Eliot GX-13 rotating anode source and mirror optics. Reflection intensities were integrated, scaled and merged using the pro-



gram XDS²³. Observations (30,048) of unique reflections (10,374) merged with an R -value of 5.4%. Data were 82% complete overall to 2.1 Å, and 82% complete with a mean $I/\sigma I = 5$ in the 2.2–2.1-Å shell. The structure was determined by molecular replacement using the AMORE²⁴ and X-PLOR²⁵ program packages. The N-terminal Syp SH2/PDGF 1009 structure¹⁵, with the peptide removed, was used as a search model. The highest and second-highest peaks in the rotation function were correct solutions for the N- and C-terminal domains of the molecule. Translation solutions for both domains were unambiguous. Rigid-body and restrained positional refinement of the molecular replacement model yielded readily interpretable electron density for the phosphopeptides and the interdomain linker, and allowed rebuilding of the C-terminal domain. The model was further refined with iterative cycles of manual rebuilding using the program O²⁶ and simulated annealing and positional refinement with X-PLOR. Ordered solvent molecules were positioned manually and with the aid of the program ARP²⁷. The final model included residues 4–153 and 166–220 of SH-PTP2, residues 1–9 and 1–10 of the peptides bound to the N- and C-terminal domains, respectively, and 348 solvent molecules. The model was refined to a crystallographic R -value of 17.2% ($R_{\text{free}} = 24.5\%$) for data with $F > 2\sigma$ and 19.9% ($R_{\text{free}} = 27.2\%$) for all data with reasonable stereochemistry (r.m.s. deviation of bonds = 0.01 Å, angles = 1.3°).

SH2 domain binding sites in the intact enzyme, we have engineered flexible linkages between 1172 and 1222 peptides of IRS-1. Individually, both IRS-1 peptides stimulate catalysis ~10–15-fold (Fig. 4b), indicating that occupancy of either domain serves a regulatory function. When linked with four sequential amino hexanoic acid (Ahx) molecules (1172Ahx₄1222), significantly higher catalytic velocities are seen at ~100-fold lower peptide concentrations. In our structure, linkages between peptides could occur in either orientation: from the C terminus of peptide N to the N terminus of peptide C (termed head-to-tail), or vice versa (tail-to-head). Functional assays show that linker orientation is critical, however. A bisphosphoryl ligand having identical composition but a tail-to-head linker orientation (1222Ahx₄1172) activates PTP less effectively than does 1172Ahx₄1222. It is possible that the PTP domain, which is not present in our structure, might interfere in the intact molecule with binding of tail-to-head linked peptides.

The length of the head-to-tail linker that yields maximal activation of the intact enzyme corresponds well with our structure of the tandem SH2 domains. For peptide 1172Ahx₄1222 (LNpYIDL₄LV(Ahx)₄LSTpYASINFQK), the distance expected for an extended chain from 1172 pTyr+5 to 1222 pTyr–2 (underlined) (41 Å) closely matches the distance between these positions on the bound 1009 peptides in the structure. Sequential shortening of the head-to-tail linker length leads to a progressive drop-off in the efficiency of catalytic activation (Fig. 4c); small increases in linker length (for example, 1172Ahx₅1222) do not. Significantly

longer linkages lead to reduced catalytic efficiency, presumably owing to entropic effects. Therefore, our structure is likely to provide an accurate model for the orientation of the domains in the active enzyme.

In contrast to IRS-1, where two phosphorylation sites appear to interact with the tandem SH2 domains, only one (Tyr 1009) of the seven documented phosphorylation sites in the PDGF receptor is thought to interact with SH-PTP2 (refs 7,8). Tyrosines 1009 and 1021 of the PDGF receptor are spaced similarly to known collinear docking sites for tandem domains (for example, ZAP-70 binding to TAM motifs¹⁷), but mutation of Tyr 1021 does not interfere in cells with SH-PTP2 binding. This result is consistent with the activation data (Fig. 4a) and structure presented here, in which peptide binding sites are widely spaced and oppositely oriented. *In vitro*, the 1009 peptide binds both SH2 domains of SH-PTP2 with equivalent affinity. Therefore, the cellular studies suggest that SH-PTP2 binds either one Tyr 1009 of a PDGF receptor monomer using one SH2 domain, or two Tyr 1009 residues in the activated receptor homodimer using both SH2 domains. The structure and activation data suggest that the latter is more likely, but our results alone cannot discriminate between the two.

With equivalent cysteines and interdomain linker lengths, it is likely that a structure of the tandem SH2 domains of SH-PTP1 would resemble that presented here. The ZAP-70 and Syk tyrosine kinases found in cells of the immune system are similar to the SH2 domain PTPs, as both contain two SH2 domains and

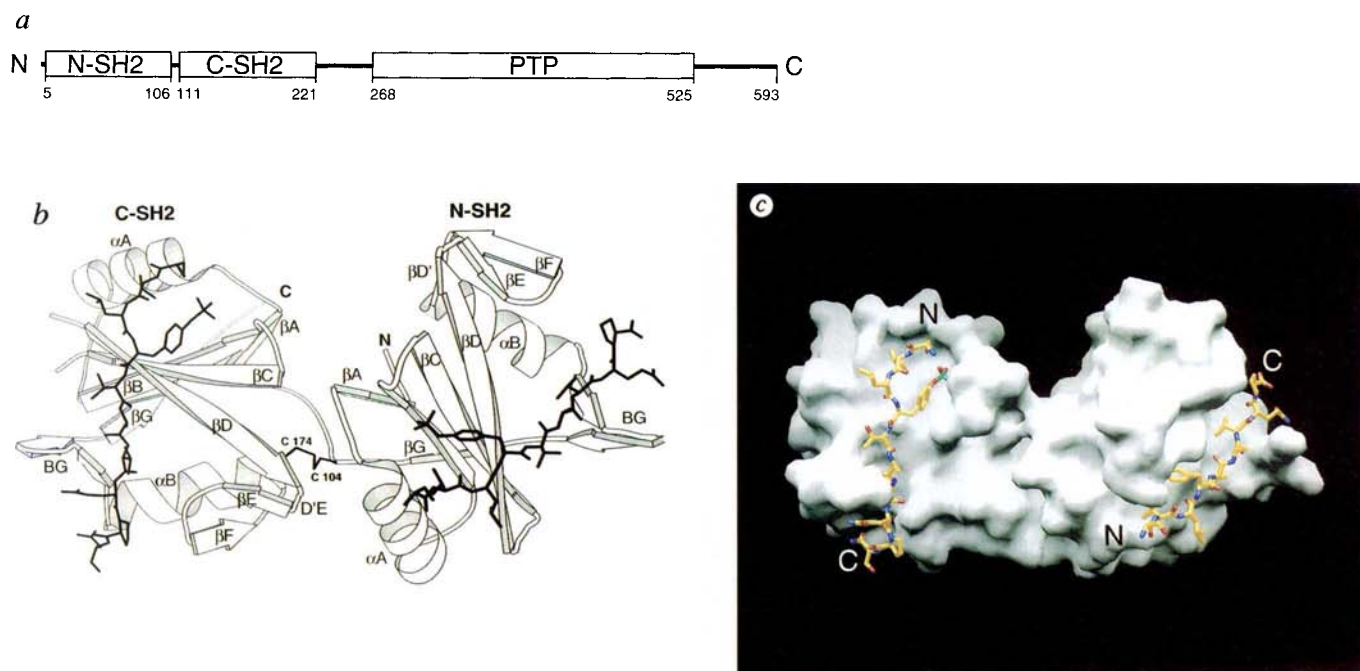


FIG. 2a, Domain architecture of SH-PTP2. b, Richardson diagram showing the secondary structure and relative orientations of domains and bound ligands in the structure of the tandem SH2 domains of SH-PTP2. Elements of secondary structure are labelled using the notation previously described for Lck²⁸. Both the N-SH2 and C-SH2 domains have short segments of antiparallel β -strand in their BG loops, a feature not found in the SH2 domains of Src-family kinases. An insertion of 10 residues in the CD loop of the C-SH2 domain relative to most other SH2 domains is disordered in this

structure. c, Molecular surface of the tandem SH2 domains, portrayed in an orientation similar to that in b. Bound phosphopeptides are shown in a stick representation. The interface between the SH2 domains creates a substantial solvent-excluded region. This structure indicates a fixed orientation of the domains relative to one another, with the phosphopeptides in widely spaced and roughly antiparallel orientations.

METHODS. MOLSCRIPT²⁹ and GRASP³⁰ were used to prepare b and c, respectively.

single C-terminal catalytic domains and show catalytic regulation by their SH2 domains¹⁸. A recent crystallographic study shows that the tandem domains of ZAP-70 in complex with a T-cell ζ -chain-derived bisphosphoryltyrosyl-based activation motif (TAM) also have a fixed orientation relative to one another¹⁷. In that case, however, the two phosphopeptide binding sites are spatially collinear, in contrast to the wide spacing and opposite orientation of binding sites in the tandem SH2 domains of SH-PTP2. In

phosphatidylinositol-3-OH kinase, which has independent catalytic (p110) and regulatory (p85) subunits, enzymatic activity is also influenced by occupancy of its two SH2 domains^{19,20}. The p85 SH2 domains are separated by ~ 190 residues, which interact directly with the p110 subunit to regulate catalysis²¹. Thus interdomain linkages are important for fixing the orientations of the SH2 domains and transmitting allosteric signals to catalytic portions of the proteins. SH2 domains often appear together with other

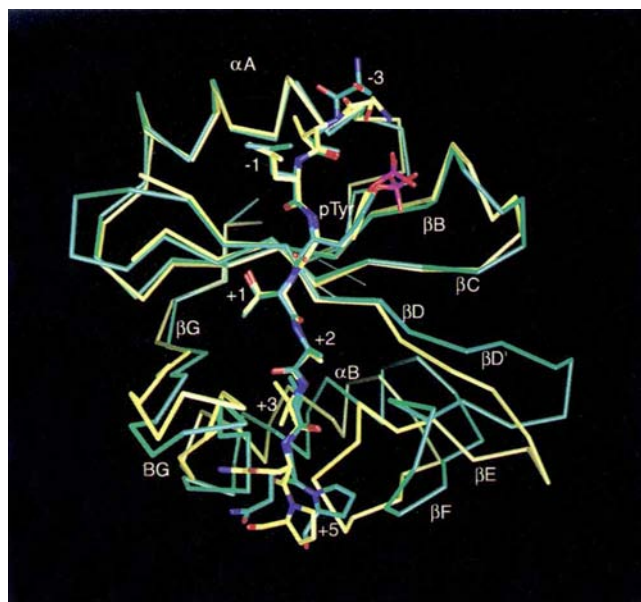
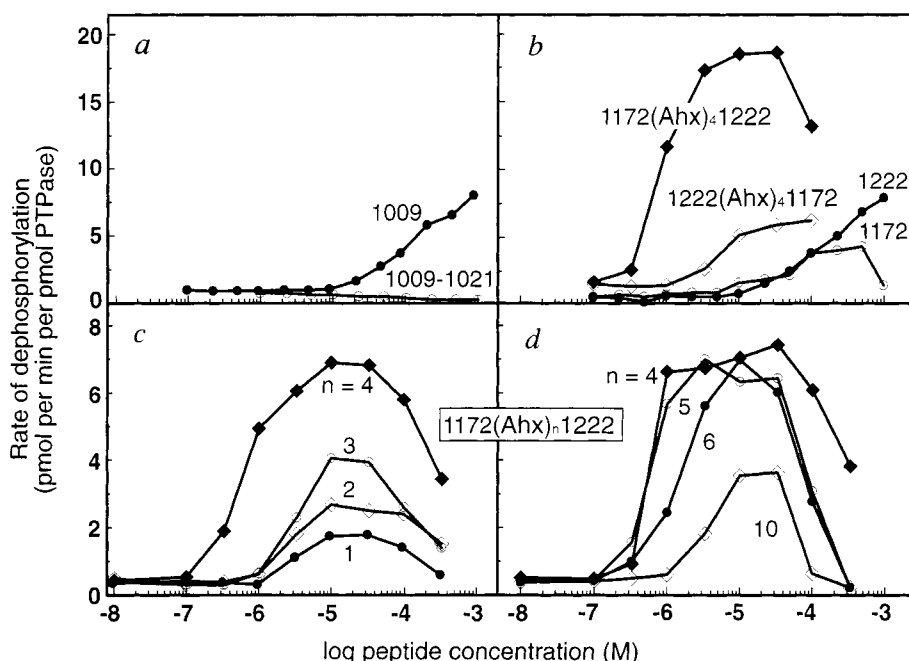


FIG. 3 Superposition of the C α traces of the N-SH2 (green) and C-SH2 (yellow) domains of SH-PTP2 and bound PDGF receptor 1009 peptides. The least-squares superposition was calculated using 44 C α atoms of the central sheet and the A helix, which superimpose with an r.m.s.d. of 0.56 Å. The structures diverge in the D', E and F strands and in the position of the B helix, which is shifted along its axis by ~ 3 Å in the N-SH2 domain. Comparisons with other domains of known structure show an altered interaction of the B helix with the central sheet in the N-SH2 domain, perhaps as a consequence of the substitution of alanine for valine in the high conserved 'FLVRES' sequence of strand β B. Despite the structural divergence between SH2 domains, the bound peptides adopt nearly identical conformations, and the mean distance between α -carbons of the seven well-ordered residues (pTyr-2 to pTyr+4) is only 0.44 Å in this superposition. The most significant divergence between peptide conformations is at phosphotyrosine itself. The phosphate group is rotated by $\sim 75^\circ$ about the C-O bond, resulting in a different network of hydrogen bonds to phosphate oxygens. Most of the residues in the pTyr pocket are conserved between the N- and C-terminal domains, but the divergence in structure may be explained by substitution of Val 148 for Thr at position β C3 in the C-SH2 domain, which eliminates a hydrogen bond.

FIG. 4 Rates of SH-PTP2-catalysed substrate dephosphorylation are influenced by SH2 domain ligands. **a**, Catalytic activation by PDGF receptor peptides 1009 (SVLPYAVQPNE) and 1009-1021 (SVLPYAVQPNEGNDNDPVIPLDPK). The 5–10-fold activation with 1009 is similar to that observed previously¹⁴; the bisphosphoryl peptide 1009-1021 inhibits rather than stimulates catalysis, presumably because the 1021 site is unable to bind the second SH2 domain and interacts instead with the catalytic domain. **b**, Activation by IRS-1 peptides 1172 (SLNPYIDLVLK) and 1222 (LSTPYASINFQK), which interact selectively with the N- and C-SH2 domains of SH-PTP2, respectively, is contrasted with that of bisphosphoryl peptides generated by covalent attachment of 1172 and 1222 with 4 amino hexanoic acid (Ahx) residues in head-to-tail (1172Ahx₄1222; LNPYIDLVL(Ahx)₄LSTPYASINFQK) versus tail-to-head (1222Ahx₄1172; LSTPYASINFQ(Ahx)₄LNPYIDLVLK) orientations. **c**, **d**, The effects of head-to-tail linker length were tested using a series of bisphosphoryl peptides differing in the number (*n*) of amino hexanoic acid residues between segments (1172Ahx_{*n*}1222, *n* = 1–10). The inhibition of PTP activity occurring at higher bisphosphoryl peptide concentrations probably follows dual occupancy of two SH2 domains by two independent bisphosphoryl peptides, leaving a free phosphotyrosine on each to interact with the PTP domain and act as a competitive substrate inhibitor. **METHODS**. Intact human SH-PTP2 protein was generated by subcloning the SH-PTP2 cDNA (provided by B. Neel, Beth Israel Hospital) into the *Nde*I site of the pET-30 vector by PCR. *E. coli* strain BL21 was transformed with the vector and protein expression was induced, cells were collected and



lysed, and proteins were isolated and purified by affinity chromatography using immobilized phosphotyrosine as described previously²⁸. After FPLC gel-filtration (Superdex-200), the purified enzyme exhibited kinetic characteristics indistinguishable from those reported previously¹³. The substrate (monophosphoryl [³²P]RCM lysozyme), peptides and SH-PTP2 enzyme were combined in a HEPES buffer, pH 7.4, and allowed to react for 5 min at 30 °C as described¹⁷. Reactions were terminated by addition of activated charcoal, and product formation was measured as [³²P] phosphate release.

regulatory or functional elements, including SH3, PTB, PH, DNA-binding and catalytic domains. In addition to direct peptide or DNA sequence recognition, spatial constraints between domains are likely to provide a much higher level of biological specificity for the intact proteins than would be achieved by the additive specificities of the isolated domains. □

Received 16 October; accepted 14 November 1995.

1. Pawson, T. *Nature* **373**, 573–580 (1995).
2. Cohen, G. B., Ren, R. & Baltimore, D. *Cell* **80**, 237–248 (1995).
3. Feng, G.-S. & Pawson, T. *Trends Genet.* **10**, 54–58 (1993).
4. Kuhne, M. R., Pawson, T., Lienhard, G. E. & Feng, G.-S. *J. biol. Chem.* **268**, 11479–11481 (1993).
5. Feng, G.-S., Hui, C.-C. & Pawson, T. *Science* **259**, 1607–1614 (1993).
6. Vogel, W., Lammers, R., Huang, J. & Ullrich, A. *Science* **259**, 1611–1614 (1993).
7. Lechleider, R. J. et al. *J. biol. Chem.* **268**, 21478–21481 (1993).
8. Kazlauskas, A., Feng, G. S., Pawson, T. & Valius, M. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6939–6942 (1993).
9. Yi, T. L., Mui, A. L.-F., Krystal, G. & Ihle, J. N. *Molec. cell. Biol.* **13**, 7577–7586 (1993).
10. Klingmüller, U., Lorenz, U., Cantley, L. C., Neel, B. G. & Lodish, H. F. *Cell* **80**, 729–738 (1995).
11. D'Ambrosio, D. et al. *Science* **268**, 293–296 (1995).
12. Doody, G. M. et al. *Science* **269**, 242–244 (1995).
13. Sugimoto, S., Wandless, T., Shoelson, S. E., Neel, B. G. & Walsh, C. T. *J. biol. Chem.* **269**, 13614–13622 (1994).

14. Pluskey, S., Wandless, T. J., Walsh, C. T. & Shoelson, S. E. *J. biol. Chem.* **270**, 2897–2900 (1995).
15. Lee, C.-H. et al. *Structure* **2**, 423–438 (1994).
16. Lechleider, R. J., Freeman, R. M. & Neel, B. G. *J. biol. Chem.* **268**, 13434–13438 (1993).
17. Hatada, M. H. et al. *Nature* **377**, 32–38 (1995).
18. Shiue, L., Zoller, M. J. & Brugge, J. S. *J. biol. Chem.* **270**, 10498–10502 (1995).
19. Backer, J. M. et al. *EMBO J.* **11**, 3469–3479 (1992).
20. Carpenter, C. L. et al. *J. biol. Chem.* **268**, 9478–9483 (1993).
21. Hu, Q., Klippel, A., Muslin, A. J., Fantl, W. J. & Williams, L. T. *Science* **268**, 100–102 (1995).
22. Case, R. D. et al. *J. biol. Chem.* **269**, 10467–10474 (1994).
23. Kabsch, W. J. *J. appl. Crystallogr.* **21**, 916–924 (1988).
24. Navaza, J. in *Molecular Replacement: Proceedings of the CCP4 Study Weekend* (ed. Dodson, E. J., Gover, S. & Wolf, W.) 87–90 (SERC, Daresbury Laboratory, Warrington, 1992).
25. Brunger, A. T. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
26. Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. *Acta Crystallogr.* **47**, 110–119 (1991).
27. Lamzin, V. S. & Wilson, K. S. *Acta crystallogr.* **D49**, 129–147 (1993).
28. Eck, M. J., Shoelson, S. E. & Harrison, S. C. *Nature* **362**, 87–91 (1993).
29. Kraulis, P. J. *Appl. Crystallogr.* **24**, 946–950 (1991).
30. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins Struct. Funct. Genet.* **11**, 281–296 (1995).

ACKNOWLEDGEMENTS. We thank B. G. Neel, R. T. Nolte and C. T. Walsh for discussions and advice. This work was supported in part by grants from the Boston Medical Foundation (M.J.E.), the Lucille P. Markey Charitable trust to Children's Hospital (M.J.E.), and the National Science Foundation (S.E.S.). S.C.H. is an investigator of the Howard Hughes Medical Institute. S.E.S. is the recipient of a Burroughs Wellcome Fund Scholar Award in Experimental Therapeutics. The Joslin Diabetes Center biochemistry facility is supported by a Diabetes & Endocrinology Research Center grant from the NIH.